

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2004, 00:29:03 ; Search time 57 Seconds
(without alignments)
2807.166 Million cell updates/sec

Title: US-09-875-519A-22
Perfect score: 8609
Sequence: 1 MPTSGPSLLLLLLHPLPA.....KQCQLGAFPTSNVFGCPN 1663

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Watch 100%
Listing first 45 summaries

Database : PIR 79:*

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8598	99.9	1663	1 C3HU	complement C3 prec
2	6825	79.3	1663	1 C3RT	complement C3 prec
3	6792.5	78.9	1666	1 C3GP	complement C3 prec
4	6749	78.4	1663	1 C3MS	complement C3 prec
5	4518.5	52.5	1852	2 I50711	complement C3 prec
6	4389.5	51.0	1851	1 C3HJ	complement C3 prec
7	4224	49.1	1642	2 I51018	cobra venom factor
8	3452.5	40.1	1620	2 I51339	complement compone
9	3082	35.8	726	2 A27602	complement C3 - ra
10	2490.5	28.9	1673	2 I50806	complement compone
11	2438.5	28.3	1230	2 T30517	complement C3-Q2 -
12	2148	25.0	1820	2 S21045	complement compone
13	2122	24.6	1683	2 T30855	complement compone
14	1840	21.4	1676	1 C5HU	complement C5 prec
15	1815	21.1	1738	1 C5MS	complement C4 prec
16	1803	20.9	1744	1 A24558	complement C4A pre
17	1733.5	20.1	1735	2 S54784	sex-limited protei
18	1732.5	20.1	1736	2 A21876	sex-limited protei
19	1515.5	17.6	1699	2 T14074	complement compone
20	1465	17.0	1617	2 T28153	complement C4 - ch
21	1465	17.0	1617	2 T28153	complement C4 - ch
22	1017	11.8	1474	1 MAHU	alpha-2-macroglobu
23	1003	11.7	1472	2 A26122	alpha-2-macroglobu
24	979	11.4	1507	2 T18544	alpha-2-macroglobu
25	967	11.2	1476	2 JC5143	alpha-1-macroglobu
26	944.5	11.0	1500	2 A42210	alpha-1-macroglobu
27	939	10.9	1477	2 A29952	alpha-1 proteinase
28	936	10.9	1487	2 S15904	alpha-1 proteinase
29	935.5	10.9	1464	2 JC5144	murinoglobulin pre

30 934.5 10.9 1495 2 S27001
31 932 10.8 1503 2 T43166
32 929.5 10.8 1473 2 A20872
33 924.5 10.7 1476 2 A41185
34 912 10.6 1482 2 S13495
35 891.5 10.4 1451 2 B41185
36 849 9.9 322 2 A32329
37 814.5 9.5 1519 2 T27828
38 805 9.4 1508 2 T27828
39 621 7.2 231 2 S27041
40 595 6.9 785 2 A29953
41 463 5.4 609 2 A41081
42 457 5.3 594 2 I48771
43 302.5 3.5 1536 2 E72310
44 288 3.3 77 2 A01259
45 249 2.9 73 2 I48284

ALIGNMENTS

RESULT 1

C3HU

complement C3 precursor [validated] - human

N:Contains: alternative complement-pathway C3/C5 convertase (EC 3.4.21.47) C3b subunit,

C:Species: Homo sapiens (man)

C:Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 09-Jul-2004

A:Accession: A94065; A37999; A92187; A27603; A23435; A45830; B45830; A01257; A01258

R:de Bruijn, M.H.L.; Fey, G.H.

Proc. Natl. Acad. Sci. U.S.A. 82, 708-712, 1985

A:Title: Human complement component C3: cDNA coding sequence and derived primary structure

A:Reference number: A94065; MUID:85140166; PMID:2579379

A:Accession: A94065

A:Molecule type: mRNA

A:Residues: 1-1663 <DEB>

A:Cross-references: UNIPROT:P01024; GB:K02765; NID:G179664; PIDN:AAA85332.1; PID:G17966

R:Viik, D.P.; Amiguet, P.; Moffat, G.J.; Fey, M.; Amiguet-Barras, F.; Wetsel, R.A.; Tack

Biochemistry 30, 1080-1085, 1991

A:Title: Structural features of the human C3 gene: intron/exon organization, transcript

A:Reference number: A37999; MUID:9113687; PMID:1703437

A:Contents: intron/exon structure of gene

A:Accession: A37999

A:Molecule type: DNA

A:Residues: 1-25 <VIK>

A:Cross-references: GB:M63423

A:Note: the authors translated the codon GGT for residue 6 as Leu, CCC for residue 7 as

R:Hugli, T.E.

J. Biol. Chem. 250, 8293-8301, 1975

A:Title: Human anaphylatoxin (C3a) from the third component of complement.

A:Reference number: A92187; MUID:76069169; PMID:1238393

A:Accession: A92187

A:Molecule type: protein

A:Residues: 672-680, 'N', 682-699, 'Q', 701-748 <HUG>

R:Doudaki, M.E.; Becherer, J.D.; Lambiris, J.D.

J. Immunol. 140, 1577-1580, 1988

A:Title: A 34-amino acid peptide of the third component of complement mediates properdin.

A:Reference number: A27603; MUID:88154452; PMID:3279119

A:Accession: A27603

A:Molecule type: protein

A:Residues: 1409-1563 <DAO>

R:Heilmann, U.; Eggertsen, G.; Engstrom, A.; Sjoquist, J.

Biochem. J. 230, 353-361, 1985

A:Title: Amino acid sequence of the trypsin-generated C3d fragment from human complement

A:Reference number: A23435; MUID:86025442; PMID:3876831

A:Accession: A23435

A:Molecule type: protein

A:Residues: 1002-1022, 'E', 1014-1303 <HEL>

A:Note: sequence corresponding to residues 1072-1100 was not determined but was taken f

R:Poznansky, M.C.; Clissold, P.M.; Lachmann, P.J.

J. Immunol. 143, 1254-1258, 1989

A:Title: The difference between human C3F and C3S results from a single amino acid chan

3.

A:Reference number: A45830; MUID:89309808; PMID:2473125

A:Accession: A45830		301	VLSRKVLDDGVONPRAEDLVGKSLVYSATVILHSGSDMVQAEBSGIPVITSPYQIHFTKT	360
A:Status: not compared with conceptual translation				
A:Molecule type: DNA		301	VLSRKVLDDGVONLRAEDLVGKSLVYSATVILHSGSDMVQAEBSGIPVITSPYQIHFTKT	360
A:Residues: 1212-1215, N', 1217-1223 <PO2>				
A>Note: this is the C3S allele		361	PKYFKPGMPFDLMVFVTPNPGSPAYRVPVAVQEDTQSLTQDGVAKLSINHPKOKPL	420
A:Accession: B45830		361	PKYFKPGMPFDLMVFVTPNPGSPAYRVPVAVQEDTQSLTQDGVAKLSINHPKOKPL	420
A:Status: not compared with conceptual translation				
A:Molecule type: DNA		421	SITVTRTKQELSEAEQATRTMCAQPYSTVGNNSNNYHLVLRTELREGETLNVEFLRMD	480
A:Residues: 1212-1223 <PO2>		421	SITVTRTKQELSEAEQATRTMCAQPYSTVGNNSNNYHLVLRTELREGETLNVEFLRMD	480
F:Doimer, K.; Sottrup-Jensen, L.				
F:FEBS Lett. 315, 85-90, 1993				
A:Title: Disulfide bridges in human complement component C3b.		481	RAHEAKIRYTYILINMKGRLLKAGQVRPGQDILVPLSLITDTPSPFLVAYYILIGA	540
A:Contents: annotation; disulfide bonds		481	RAHEAKIRYTYILINMKGRLLKAGQVRPGQDILVPLSLITDTPSPFLVAYYILIGA	540
C:Comment: The sequence shown is the C3 fast (C3p) allele, which is found mainly in Caucasian alternative complement pathways, releases the C3a anaphylatoxin from the amino end of the native-complement-pathway C3/C5 convertase.		541	SCOREWADSVWVDVKSCVGLVVKSGQSDROPVPGQQTWLIKEGDHGARVVLVAVDK	600
C:Comment: C3a anaphylatoxin is a vasoactive peptide and a mediator of inflammation.		541	SCOREWADSVWVDVKSCVGLVVKSGQSDROPVPGQQTWLIKEGDHGARVVLVAVDK	600
C:Comment: C3b, with its highly reactive thiol group, binds to the surface of foreign particles classical-complement-pathway C3/C5 convertase. The activity of C3b is regulated by protein C.		601	GVFVLNKKXKLTQSKINDVVEKADIGCTPGSGKDYAGVFS DAGLFTTSSGQQTAAEAL	660
C:Genetics:		601	GVFVLNKKXKLTQSKINDVVEKADIGCTPGSGKDYAGVFS DAGLFTTSSGQQTAAEAL	660
A:Gene: GDB:C3				
A:Cross-references: GDB:119044; OMIM:120700		661	QCQPAARRRRSVQITKMDKVKYKELKCCEDGWRPNPFRFSCQRTTRFISLGEAC	720
A:Map position: 19p13.3-19p13.3		661	QCQPAARRRRSVQITKMDKVKYKELKCCEDGWRPNPFRFSCQRTTRFISLGEAC	720
A>Note: contains 41 exons				
C:Superfamily: alpha-2-macroglobulin		721	KKVFLDCCNYITELRRQHARASHGLARSLNDEDIIAEENIVSRSEPPESLWVVEDLKE	780
C:Keywords: acute phase; complement alternate pathway; complement pathway; glycoprotein; F1-22/Domain: signal sequence #status predicted <SIG>		721	KKVFLDCCNYITELRRQHARASHGLARSLNDEDIIAEENIVSRSEPPESLWVVEDLKE	780
F:23-667/Product: complement C3 and C3b beta chain #status predicted <C3BB>				
F:23-667,672-1663/Product: complement C3 #status predicted <CC3>				
F:23-667,749-1663/Product: C3b #status predicted <C3B>				
F:672-1663/Product: complement C3 alpha chain #status predicted <CC3A>		781	PPNGISTKLMNIFLKDSITTWELAVSMSSDKKICVADPEVTVMDOFFIDILRLPYSVV	840
F:672-749/Product: C3a anaphylatoxin #status predicted <C3T>		781	PPNGISTKLMNIFLKDSITTWELAVSMSSDKKICVADPEVTVMDOFFIDILRLPYSVV	840
F:749-1663/Product: C3b alpha' chain #status predicted <C3BA>				
F:946-1303/Product: C3dk fragment #status predicted <CDK>				
F:955-1303/Product: C3dg fragment #status predicted <CDG>		841	RNEQVEIRAVLYNYPQNOELKVRVELLHNPFCSLATTKRHHQQTITIPKSSLSVPYVI	900
F:955-1001/Product: C3g fragment #status predicted <C3G>		841	RNEQVEIRAVLYNYPQNOELKVRVELLHNPFCSLATTKRHHQQTITIPKSSLSVPYVI	900
F:1002-1303/Product: C3d fragment #status experimental <C3D>				
F:1424-1457/Region: properdin binding				
F:85-939/Binding site: carbohydrate (Asn) (covalent) #status experimental		901	VPLKTLGLEVEVKAAYVHHFTISDGVKSLKVPPEGIRMKTVAVRTLDPERLREGVQKE	960
F:559-816/Binding site: carbohydrate (Asn) (covalent) #status experimental		901	VPLKTLGLEVEVKAAYVHHFTISDGVKSLKVPPEGIRMKTVAVRTLDPERLREGVQKE	960
F:748-749/Cleavage site: Arg-Ser (C3 convertase) #status predicted				
F:954-955/Cleavage site: Arg-Glu (complement factor I) #status predicted		961	DIPPADLSQVPTDSETRILLQGTTPVAQMTEDAVDAERLKHILVTPSGCGEQNMIGTWP	1020
F:1010-1013/Cross-link: thiolester (Cys-Gln) #status experimental		961	DIPPADLSQVPTDSETRILLQGTTPVAQMTEDAVDAERLKHILVTPSGCGEQNMIGTWP	1020
F:1303-1304/Cleavage site: Arg-Ser (complement factor I) #status predicted				
F:1320-1321/Cleavage site: Arg-Ser (complement factor I) #status predicted				
F:1617/Binding site: carbohydrate (Asn) (covalent) #status predicted		1021	TVIAVHYLDETEQWEKFGLEKRGQALELIKKGYTQQLAFROPSSAPAAFAVVRAPSTWLT	1080
Query Match				
Best Local Similarity 99.9%; Score 8598; DB 1; Length 1663;				
Matches 1661; Conservative 1; Mismatches 1; Indels 0; Gaps 0;				
QY	1	MGPTSGPSLLLLTHLPLALGSPMYSITPNILRSEETWVLEAHDAQDVPVTVTVH	60	
DB	1	MGPTSGPSLLLLTHLPLALGSPMYSITPNILRSEETWVLEAHDAQDVPVTVTVH	60	
QY	61	DFPGKLVLSSEKTVLTATNMGVNTFTIPANREFKSEKGNKFNFTVQATFGVQVEKV	120	
DB	61	DFPGKLVLSSEKTVLTATNMGVNTFTIPANREFKSEKGNKFNFTVQATFGVQVEKV	120	
QY	121	VLVLSQSGYLFITQDNTIYTPGSTVLYRIFTVNHKLLPVGRVVMVNIENPEGIPVKQDSL	180	
DB	121	VLVLSQSGYLFITQDNTIYTPGSTVLYRIFTVNHKLLPVGRVVMVNIENPEGIPVKQDSL	180	
QY	181	SSQNQLGVLPDWDPELVNMGQWIRAYENSPQVSTFEFVKYVLPSEFVIVTEPE	240	
DB	181	SSQNQLGVLPDWDPELVNMGQWIRAYENSPQVSTFEFVKYVLPSEFVIVTEPE	240	
QY	241	KFYIYINXKGLVTTITARELYGKVGEGTAFVIFGQDGEQRISLPESLKRIPIEDGSGEV	300	
DB	241	KFYIYINXKGLVTTITARELYGKVGEGTAFVIFGQDGEQRISLPESLKRIPIEDGSGEV	300	

QY 1381 KNTMILEICTRYGDDATWSILDISMTGTFAPDITDLDKOLANGVDVYISKYELDKAFSD 1440
 Db 1381 KNTMILEICTRYGDDATWSILDISMTGTFAPDITDLDKOLANGVDVYISKYELDKAFSD 1440
 QY 1441 RNTLIYLDKVSHEDDCLAFKQVQFNVELIQGAVKYVAYYNLEESCTRFYHPEKEDG 1500
 Db 1441 RNTLIYLDKVSHEDDCLAFKQVQFNVELIQGAVKYVAYYNLEESCTRFYHPEKEDG 1500
 QY 1501 KLNKLGRDELCAEENCFTQKSDDKVTLEERLDKACEGVDYVYKTRLVKQLSNDPDE 1560
 Db 1501 KLNKLGRDELCAEENCFTQKSDDKVTLEERLDKACEGVDYVYKTRLVKQLSNDPDE 1560
 QY 1561 YIMAIQTIKSGSDEVQVQQTFTFISPIKREALKLEEKHYLMWGLSDFWGEKPNLSY 1620
 Db 1561 YIMAIQTIKSGSDEVQVQQTFTFISPIKREALKLEEKHYLMWGLSDFWGEKPNLSY 1620
 QY 1621 IIGKDTWVHEWPEDECOBENKQKODYGATESWVFGCPN 1663
 Db 1621 IIGKDTWVHEWPEDECOBENKQKODYGATESWVFGCPN 1663
 RESULT 2
 C3RT
 Complement C3 precursor - rat
 N:Alternate names: 37K phospholipase A2 inhibitory protein
 N:Contains: alternative-complement-pathway C3/C5 convertase (EC 3.4.21.47) C3b subunit;
 C:Species: Rattus norvegicus (Norway rat)
 C:date: 04-Dec-1992 #sequence revision 07-Oct-1994 #text change 09-Jul-2004
 C:title: Estrogen regulation of tissue-specific expression of complement C3.
 R:Mišumi, Y.; Sohma, M.; Ikebara, Y.
 Nucleic Acids Res. 18, 2178, 1990
 A:Title: Nucleotide and deduced amino acid sequence of rat complement C3.
 A:Reference number: S15764; MUID:190245672; PMID:2336397
 A:Accession: S15764
 A:Molecule type: mRNA
 A:Residues: 1-1663 <MIS>
 A:Cross-references: UNIPROT:P01026; EMBL:X52477; NID:G56953; PIDN:CRAA36716.1; PID:G56954
 R:Sunstom, S.A.; Komm, B.S.; Ponce-de-Leon, H.; Yi, Z.; Teuscher, C.; Lyttle, C.R.
 J. Biol. Chem. 264, 16941-16947, 1989
 A:Title: Estrogen regulation of tissue-specific expression of complement C3.
 A:Reference number: A54562; MUID:89380332; PMID:2674144
 A:Accession: A54562
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 'P', 1316-1595 <SUN>
 A:Cross-references: GB:M29866; NID:9203200; PIDN:AAA40837.1; PID:G554423
 R:Jacobs, J.W.; Rubin, J.S.; Hugli, T.E.; Bogardt, R.A.; Mariz, I.K.; Daniels, J.S.; Dau
 Biochemistry 17, 5031-5038, 1978
 A:Title: Purification, characterization, and amino acid sequence of rat anaphylatoxin (C3a)
 A:Reference number: A01260; MUID:79062262; PMID:309768
 A:Accession: A01260
 A:Molecule type: protein
 A:Residues: 671-703, 'K', 705-720, 'KL', 723-748 <JAC>
 A:Note: three disulfide bonds are present
 R:Suwa, Y.; Kudo, I.; Imaizumi, A.; Okada, M.; Kamimura, T.; Suzuki, Y.; Chang, H.W.; Ha
 Proc. Natl. Acad. Sci. U.S.A. 87, 2395-2399, 1990
 A:Title: Proteinaceous inhibitors of phospholipase A-2 purified from inflammatory sites
 A:Reference number: A35979; MUID:90207203; PMID:2320562
 A:Accession: B35979
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 'X', 998-1005 <SUN>
 A:Accession: A35979
 A:Molecule type: protein
 A:Residues: 'X', 961-962, 'P', 964-969 <SU2>
 R:Nakagawa, H.; Komoriya, N.
 Biochem. Biophys. Res. Commun. 194, 1181-1187, 1993
 A:Title: Complement component C3-derived neutrophil chemotactic factors purified from ex
 A:Reference number: PNO566; MUID:93356786; PMID:8352775
 A:Accession: PNO567
 A:Molecule type: protein
 A:Residues: 568-592 <NAK>
 A:Note: amino end of a C3-derived peptide designated exudate neutrophil chemotactic fact

A:Accession: PNO566
 A:Molecule type: protein
 A:Residues: 671-687 <NA2>
 A:Note: amino end of peptide designated neutrophil chemotactic factor 1 and probably id
 R:Kuivaneen, P.C.; Capulong, R.B.; Harkins, R.N.; Desombre, E.R.
 Biochem. Biophys. Res. Commun. 158, 898-905, 1989
 A:Title: The estrogen-responsive 110K and 74K rat uterine secretory proteins are structu
 A:Reference number: A32281; MUID:89149812; PMID:2645873
 A:Accession: A32281
 A:Molecule type: protein
 A:Residues: 25-41 <KUI>
 A:Experimental source: 17beta-estradiol-stimulated uterus of immature rat
 A:Note: the authors treat this 74K uterine secretory protein, identical as far as sequen
 C:Comment: Complement C3 contains two chains, formed by removal of four residues and 11
 alternative-complement-pathway C3/C5 convertase.
 C:Comment: C3a anaphylatoxin is a vasoactive peptide and a mediator of inflammation.
 C:Comment: C3b, with its highly reactive thiol group, binds to the surface of foreign p
 e classical-complement-pathway C3/C5 convertase. The activity of C3b is regulated by pr
 C:Superfamily: alpha-2-macroglobulin
 C:Keywords: acute phase; chemotaxis; complement alternate pathway; complement pathway;
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-666/Product: complement C3 and C3b beta chain #status predicted <C3BB>
 F:25-666, 671-1663/Product: complement C3 #status predicted <CC3>
 F:25-666, 749-1663/Product: complement C3b #status predicted <CC3>
 F:671-1663/Product: complement C3 alpha chain #status predicted <CC3A>
 F:671-748/Product: C3a anaphylatoxin #status experimental <C3T>
 F:749-1663/Product: complement C3b alpha' chain #status predicted <C3BA>
 F:946-1303/Product: C3dk fragment #status predicted <CDK>
 F:1002-1303/Product: C3d fragment #status predicted <C3D>
 F:1424-1457/Region: properdin binding
 F:558-816, 626-661, 693-720, 694-727, 707-728, 873-1513, 1101-1158, 1358-1489, 1389-1458, 1506-1
 F:748-749/Cleavage site: Arg-Ser (C3 convertase) #status predicted
 F:939, 1617/Binding site: Carbohydrate (Asn) (covalent) #status predicted
 F:1010-1013/Cross-link: thiolester (Cys-Gln) #status predicted
 F:1303-1304/Cleavage site: Arg-Ser (complement factor I) #status predicted
 F:1320-1321/Cleavage site: Arg-Ser (complement factor I) #status predicted
 Query Match 79.3%; Score 6825; DB 1; Length 1663;
 Best Local Similarity 78.1%; Pred. No. 0;
 Matches 1301; Conservative 171; Mismatches 188; Indels 6; Gaps 5;
 QY 1 MGPTSGPSL--LILLTLPLALGSPMSYIIIPNIRLESEETMVLAEHDAQDVPVTV 58
 Db 1 MGPTSGQLLVLLILLASSILLALGSPMSYIIIPNIRLESEETFILEHDAQDVPVTV 60
 QY 59 VHDFFPKKLVLSSEKTVLTPATNMGNTVFTIPANREFKSEKGRNKFVTVQATFGQVYE 118
 Db 61 VQDFL-KKQVLTSEKTVLTGATGHLNRVFIKIPASKEFNADKG-HKYVTVWANFGATVVE 118
 QY 119 KVLVSLQSGYLFIOQDKTIYTPGSTVLYRIFTVNHKLLPVGRTVMMNIENPEGIPVKOD 178
 Db 119 KAVLVVSFQSGYLFIOQDKTIYTPGSTVLYRIFTVNNLLPVGRTVMMNIENPEGIPVKOD 178
 QY 179 SLSSNQGLVPLSDWIDPELVNMGQKIRAYVNSPQQVFSTEFVKEYVLSFVFEV 238
 Db 179 ILSSHNQYGLPLSNWNPILVNNMGQKIRAFYHAPKQTFSAEFVKEYVLSFVFEV 238
 QY 239 TEKFYIYNEKGLVITITARELYGKVGTAIVIGIODGEORISLPESLKXIPEDSGG 298
 Db 239 TEKFYIHGPKGLEVISITARFLYKQNVGTAVIFGVQDQDKKISLAUSLTVLVEDSGG 298
 QY 299 EVVLGRKVLIDGVQNPRAEDLVGKSLYSATVILHSGSDMWQAERSGIPVTSPIQIHFT 358
 Db 299 EAVLSRKVLMDGVRSPSEALVGKSLYSVTVILHSGSDMWQAERSGIPVTSPIQIHFT 358
 QY 359 KTPKYFKCMPDLAVFTNPDGSPAYRVPVAVQGEDTVQSLTQDGVAKLSINTHPQSK 418
 Db 359 KTPKFKPAMPFDLMVFTNPDGSPARVPVAVQGEDTVQSLTQDGVAKLSINTHPQSK 417
 QY 419 PLISITVTKKQELSEAEQATRMQALPVSTVGNSSNNYLHLSVLRLPGETLNNVFLLR 478

Db 418 PLTITVSTKKEGIPDARQARTWQAPYSTHNSNNYLRLSVSRVLPKPGDNLNWFHLR 477
QY 479 MDRAEAKRYTYTILMNGRLLKAGQVREPGQDLVVLPLSITTDIPSPRLVAVYTLI 538
Db 478 TDAGQBAKRYTYTILMNGRLLKAGQVREPGQDLVVLPLSITTDIPSPRLVAVYTLI 537
QY 539 GASGQREVVDVSVVVDKDCVSGSLVYKSGQSDQPVPGQDMTKLKGHDHARVILVAV 598
Db 538 GANGQREVVDVSVVVDKDCVSGSLVYKSGQSDQPVPGQDMTKLKGHDHARVILVAV 597
QY 599 DKGVPVFNKKNKLTQSKINDVVEKADIGCTPGSGKDVAGVPSDAGLFTSSGGQOQA 658
Db 598 DKGVPVFNKKNKLTQSKINDVVEKADIGCTPGSGKDVAGVPSDAGLFTSSGGQOQA 657
QY 659 ELQCPQAPARRRSVOLTETKRVMDKVGY-PKELRKCCEDGMREMPFSPQRRTFISLG 717
Db 658 DPECAKAPARRRSVOLTETKRVMDKVGY-PKELRKCCEDGMREMPFSPQRRTFISLG 717
QY 718 EACKKVFELDCNVITELRQAHASHGLGARSNLDDEIIAENIVSRSEPPESLWNVED 777
Db 718 EACKKVFELDCNVITELRQAHASHGLGARSNLDDEIIAENIVSRSEPPESLWNVED 777
QY 778 LKPPKNGISTKLMIIFLKDSTTWELLAVNSDKKICVADPEVTVMDQFFDILRLPY 837
Db 778 LKPPKNGISTKLMIIFLKDSTTWELLAVNSDKKICVADPEVTVMDQFFDILRLPY 837
QY 838 SVVNEQVIRAVLYNRYNQELKVRVLLHNPAPCSLATTKRHOCTITIPPKSSLSVP 897
Db 838 SVVNEQVIRAVLYNRYNQELKVRVLLHNPAPCSLATTKRHOCTITIPPKSSLSVP 897
QY 898 YVIVPLKGTQVEVKAAYHHFISDGVKRSKVVPEGRVNVKTVAVRTLDPERLREGV 957
Db 898 YVIVPLKGTQVEVKAAYHHFISDGVKRSKVVPEGRVNVKTVAVRTLDPERLREGV 957
QY 958 QKEDIIPADLSQVDPDTESETILLQGTFFVQMTEDAVDAERLKHLLVTPSGCGEQNMIG 1017
Db 958 QKEDIIPADLSQVDPDTESETILLQGTFFVQMTEDAVDAERLKHLLVTPSGCGEQNMIG 1017
QY 1018 MPTVTIANYHLDTSQWEKFGLEKQGALELIKGYTQQLAFQPIISAYAAFNRRPSTW 1077
Db 1018 MPTVTIANYHLDTSQWEKFGLEKQGALELIKGYTQQLAFQPIISAYAAFNRRPSTW 1077
QY 1078 LTAIVVVKVFLANLIIADSVLCGAVKWLILEKQKDPGVQEDAPVTHQEMIGGLRNN 1137
Db 1078 LTAIVVVKVFLANLIIADSVLCGAVKWLILEKQKDPGVQEDAPVTHQEMIGGLRNN 1137
QY 1138 EKDMALTAVFLISLOEAKDICEQVNSLPGSITKAGDFLEANYMNLORSYTVTAIAGYALA 1197
Db 1138 EADVSLTAVFLALQEARDCIGQVNSLPGSITKAGDFLEANYMNLORSYTVTAIAGYALA 1197
QY 1198 QMRLKGLPLNKFLTTAKDKNWDPPGKQVNSLPGSITKAGDFLEANYMNLORSYTVTAIAGYALA 1257
Db 1198 QMRLKGLPLNKFLTTAKDKNWDPPGKQVNSLPGSITKAGDFLEANYMNLORSYTVTAIAGYALA 1257
QY 1258 BORYYGGVSGTQATFWFQALQYQKADPHQELNLDVSLQPSRSKXTHRIHMSAS 1317
Db 1258 BORYYGGVSGTQATFWFQALQYQKADPHQELNLDVSLQPSRSKXTHRIHMSAS 1317
QY 1318 LRSSETKNEGTFTVTEGKGQGTLSVVTMTHAKADOLTCNKFDLVTKPAPETKRP 1377
Db 1318 LRSSETKNEGTFTVTEGKGQGTLSVVTMTHAKADOLTCNKFDLVTKPAPETKRP 1377
QY 1378 QAKVNTMILICTRVGGDDATMSILDSMTGTAPDITDQKQANGVDRYISKYELDKA 1437
Db 1378 QAKVNTMILICTRVGGDDATMSILDSMTGTAPDITDQKQANGVDRYISKYELDKA 1437
QY 1438 FSDRNTLIYLDKVSHEDDCLAFKVOYFNVELIQGAVKVAAYNLESCTFRFVHPEK 1497
Db 1438 FSDRNTLIYLDKVSHEDDCLAFKVOYFNVELIQGAVKVAAYNLESCTFRFVHPEK 1497
QY 1498 EDGLKNTLRCDELRCABENCFIQSDDKVLEBLRLDCAKCPGVYDYVYKTLRVKQLSND 1557
Db 1498 EDGLKNTLRCDELRCABENCFIQSDDKVLEBLRLDCAKCPGVYDYVYKTLRVKQLSND 1557

Db 1498 DGMLSKLCHNEMCRCAEENCFMHQSQDQVSLNRLDKACEPGVDVYVYKTLTIELSD 1557
QY 1558 FDEYIMATBOPIKSGSDEVQGOORTFISPIKCREALKLEKHYLMWGLSGDFWGEKEN 1617
Db 1558 FDEYIMATBOPIKSGSDEVQGOORTFISPIKCREALKLEKHYLMWGLSGDFWGEKEN 1617
QY 1618 LSYIIGKDTWVEHWPBEDEQDEENKQKQCDLGAFTSMVFGCPN 1663
Db 1618 LSYIIGKDTWVEHWPBEDEQDEENKQKQCDLGAFTSMVFGCPN 1663
RESULT 3
C3GP
Complement C3 precursor - guinea pig
NContains: alternative-complement-pathway C3/C5 convertase (EC 3.4.21.47) C3b subunit;
C.Species: Cavia porcellus (guinea pig)
C.Date: 07-Feb-1992 #sequence_revision 07-Oct-1994 #text change 09-Jul-2004
C.Accession: A37156; S03375; A20342; D20342; C20342; A31222
J.Auerbach, H.S.; Burger, R.; Dodds, A.; Colten, H.R.
J. Clin. Invest. 96, 96-106, 1990
A.Title: Molecular basis of complement C3 deficiency in guinea pigs.
A.Reference number: A37156; MUID:90307998; PMID:1973176
A.Accession: A37156
A.Molecule type: mRNA
A.Residues: 1-1666 <AUE>
A.Cross-references: UNIPROT:P12387; GB:M34054; NID:g191262; PIDN:AAA37038.1; PID:g305335
R.Gerard, N.P.; Lively, M.O.; Gerard, C.
Protein Seq. Data Anal. 1, 473-478, 1988
A.Title: Amino acid sequence of guinea pig C3a anaphylatoxin.
A.Reference number: S03375; MUID:89113342; PMID:3064079
A.Accession: S03375
A.Molecule type: Protein
A.Residues: 676-730, 'N', 732-752 <GER>
A.Experimental source: Complement-activated guinea pig serum
A.Note: Form isolated is inactive C3a anaphylatoxin and is missing the carboxyl-terminal
R.Thomas, M.L.; Tack, B.F.
Biochemistry 22, 942-947, 1983
A.Title: Identification and alignment of a thiol ester site in the third component of guinea pig serum.
A.Reference number: A90479; MUID:83178889; PMID:6838833
A.Accession: A20342
A.Molecule type: Protein
A.Residues: 676-687 <TH1>
A.Accession: D20342
A.Molecule type: Protein
A.Residues: 993-1012, 1014-1017, 'E', 1019-1030, 'Y' <TH2>
J. Biol. Chem. 256, 12617-12619, 1981
A.Title: NH2-terminal structure and cleavage of guinea pig pro-C3, the precursor of the classical complement pathway C3/C5 convertase.
A.Reference number: A20342; MUID:82075767; PMID:6458605
A.Accession: C20342
A.Molecule type: Protein
A.Residues: 23-38 <GOL>
C.Comment: Complement C3 contains two chains, formed by removal of four residues and lin
alternative complement pathways, releases the C3a anaphylatoxin from the amino end of t
native-complement-pathway C3/C5 convertase.
C.Comment: C3a anaphylatoxin is a vasoactive peptide and a mediator of inflammation.
C.Comment: C3b, with its highly reactive thiol group, binds to the surface of foreign pa
e classical complement pathway C3/C5 convertase. The activity of C3b is regulated by pr
C.Comment: The major site of synthesis of this plasma protein is the liver.
C.Superfamily: alpha-2-macroglobulin
C.Keywords: acute phase; complement alternate pathway; complement pathway; glycoprotein;
F.1-22/Domain: signal sequence #status predicted <SIG>
F.23-671/Product: complement C3 and C3b beta chain #status predicted <C3BB>
F.23-671, 676-1666/Product: complement C3 #status predicted <C3>
F.23-671, 754-1666/Product: complement C3b #status predicted <C3B>
F.676-1666/Product: complement C3 alpha chain #status predicted <CC3>
F.676-753/Product: C3a anaphylatoxin #status predicted <C3T>
F.754-1666/Product: complement C3b alpha' chain #status predicted <C3BA>
F.951-1308/Product: C3dk fragment #status predicted <CDK>
F.1007-1308/Product: C3dk fragment #status predicted <CDK>
F.1429-1461/Region: properdin binding
F.551-821, 830-866, 698-732, 712-733, 878-1517, 1106-1163, 1363-1493, 1394-1462, 1510-1
F.753-754/Cleavage site: Arg-Ser (C3 convertase) #status predicted

F;944,1620/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;1015-1018/Cross-link: thioester (Cys-Gln) #status experimental
 F;1308-1309/Cleavage site: Arg-Ser (complement factor I) #status predicted
 F;1325-1326/Cleavage site: Arg-Ser (complement factor I) #status predicted

Query Match: 78.9%; Score 6792.5; DB 1; Length 1666;
 Best Local Similarity 77.7%; Pred. No. 0;
 Matches 1297; Conservative 180; Mismatches 182; Indels 11; Gaps 7;

```

QY 1 MGPTSGPSLLLLTHLPALGSPMYSIITPNILRLESEETVLENDHQDQVPTVTVH 60
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1 MGPAAGPSLLLLLASVALGDPMSYIITPNILRLENEETVLEAHEVQGDIPVTVH 60
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 DFGKGLVLSSEKTVLTPATNHNVTFTIPANREPSEKGRNKFVTVQATFGTQVEKV 120
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 DPAKKNVLSSEKTVLTSATGLTGTIKIPASKEFKDKGR-KLVVQAAGGGTQLEKV 119
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 VLVSQSGYLFIQTDKTIITPGSTVLYRIFTVNNKLLPVGRVTVMNENPEGIPVKQSL 180
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 120 VLVSQSGYLFIQTDKTIITPGSTVLYRIFTVDSLLPVGRITIIVTIETPDGIPKRDTL 179
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 SSQNLGLVPLSDIPELVNMGOWKIRAYENSPOQVSTEPSEVEYVLPSEVIVBTE 240
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 180 SSNNQHGILPLSNIFELVNMGOWKIQAENSPKQVSAEVEYVLPSEVIVBTE 239
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 241 KFYIYINKEGLEVTITARELYKKVEGTAFVIFIGDGEQRISLPSERKRIPIEDSGEV 300
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 240 KFYIDDDPKGLENIITARELYGNVDGTAFVIFGVQDQQRISLAQSLTRVIVIEDSGEV 299
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 301 VLSRKVLDDGVNPRADLVGKSLVSVATVILHSGSDMVQAERSGIPVTSPIQHFTKT 360
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 300 VLSRQVLDDGVNPRADLVGKSLVSVATVILHSGSDMVQAERSGIPVTSPIQHFTKT 359
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 361 PKYFKPMFPDLMVFNTPDGPAPVAVVQGEDTVQSLTQDGVAKLSINTHPSKPL 420
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 360 PKYFKPAMPEIMVLTNPDGSPAPHPVVTQGSN-VQSLTQADGVARLSINTPNRQPL 418
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 421 SITVTRFKQELSEAGQATETMALPYSITVGNMNYLHLSVURTELPRGETLNVNPLRMD 480
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 419 SVTVQTKGGIPDARQAINMQALPITMYNENNYLHLSMFRTELKPGETINVNPHLRSD 478
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 481 RAHEAKIRYTYLIMNKGRLKAGRVREPGQDLVPLSITTDPIPSFRLVAVYTLIGA 540
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 479 PNOEAKIRYTYLIMNKGRLKAGRVREPGQDLVPLSITTDPIPSFRLVAVYTLIGA 538
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 541 SCOREWADSVNVDKSCVGLVYKSGQSED-----RQVPFGQMTUKIGDGHARVVL 595
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 539 SAQREVADVSWADVRDSCVGLVYKSGGKDGQDKRQQLPFRQOMTLRIEIGNQARVGL 598
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 596 VAVDKGVFVINKKNTOSKIWDVVEKADIGCTPGSGKDYAGVPSDAGLTFTSS-SGOOT 654
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 599 VAVDKGVFVINKKNTOSKIWDVVEKADIGCTPGSGKDYAGVPSDAGLTFTSS-SGOOT 658
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 655 AQRAELQCPQPAARRRRSVQLTETKMDKVGKY-PKELAKCEDGMRNPMRFSQORTRF 713
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 659 AQREGLDQCPKPAARRRRSVQLTETKMDKVGKY-PKELAKCEDGMRNPMRFSQORTRF 718
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 714 ISLGEACKKVFLLDCNYYITELRQHRASHGLARSNLDEDIIAENIVSSETPESWLW 773
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 719 VSLGEACVKAFLDCCCTYWAQRLQRRRQNLGLARSNLDEDIIAENIVSSETPESWLW 778
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 774 NVEDIKEPPKXGISLKMNIIFKDSITTWELAVSMDSKKGICVADPFVTVMDFFIDL 833
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 779 TIEELKEPERNGISTKTNWIFLKDSTTWELAVSMDSKKGICVADPFVTVMDFFIDL 838
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 834 RLPYSVWNEQVEIRAVLYNQQLKVRVELLHNPAFCSLATKRRHOOTIIPKSS 893
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 839 RLPYSVWNEQVEIRAVLYNQQLKVRVELLHNPAFCSLATKRRHOOTIIPKSS 898
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 894 LSVPTVYVPLKGLQVEVKAAYVHHFTISDGVKRSKLVVPEGIRMNKTVAVRTLDPRLG 953
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 899 VAVPYVLPKLGQVEVKAAYVNYFISDGVKRSKLVVPEGIRMNKTVAVRTLDPRLG 958
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 4

C3MS

complement C3 precursor - mouse

N;Contains: alternative-complement-pathway C3/C5 convertase (EC 3.4.21.47) C3b subunit;
 C;Species: Mus musculus (house mouse)

C;Date: 30-Jun-1988 #sequence revision 30-Jun-1988 #text change 09-Jul-2004

C;Accession: A92459; B92459; A92460; A93938; A21898; A54561; S16369; S16189; I49563; I4

R;Lundwall, A.; Wetzel, R.A.; Domdey, H.; Tack, B.F.; Fey, G.H.

J. Biol. Chem. 259, 13851-13856, 1984

A;Title: Structure of murine complement component C3: I. Nucleotide sequence of cloned

A;Reference number: A92459; MUID:85054818; PMID:6548745

A;Accession: A92459

A;Molecule type: mRNA

A;Residues: 1-724 <LU1>

A;Cross-references: UNIPROT:P01027

A;Accession: B92459

A;Molecule type: DNA

A;Residues: 1-124 <LU2>

R;Wetzel, R.A.; Lundwall, A.; Davidson, F.; Gibson, T.; Tack, B.F.; Fey, G.H.

J. Biol. Chem. 259, 13857-13862, 1984

A;Title: Structure of murine complement component C3: II. Nucleotide sequence of cloned

A;Reference number: A92460; MUID:85054819; PMID:6094532

A;Accession: A92460

A;Molecule type: mRNA

A;Residues: 671-1663 <WPT>
R;Domdey, H.; Wiebauer, K.; Kazmaier, M.; Muller, V.; Odink, K.; Fey, G.
Proc. Natl. Acad. Sci. U.S.A. 79, 7619-7623, 1982
A;Title: Characterization of the mRNA and cloned cDNA specifying the third component of
A;Reference number: A93938; MUID:83117730; PMID:6961437
A;Contents: C3a
A;Accession: A93938
A;Molecule type: mRNA
A;Residues: 671-748 <DOM>
R;Sottrup-Jensen, L.; Stepanik, T.M.; Kristensen, T.; Lonblad, P.B.; Jones, C.M.; Wierzb
Proc. Natl. Acad. Sci. U.S.A. 82, 9-13, 1985
A;Title: Common evolutionary origin of alpha2-macroglobulin and complement components C3
A;Reference number: A21898; MUID:85113177; PMID:2578664
A;Accession: A21898
A;Molecule type: mRNA
A;Residues: 25-1663 <SOT>
R;Hamada, J.; Cavanaugh, P.G.; Mikli, K.; Nicolson, G.L.
Cancer Res. 53, 4418-4423, 1993
A;Title: A paracrine migration-stimulating factor for metastatic tumor cells secreted by
A;Reference number: A54561; MUID:93373334; PMID:8364938
A;Accession: A54561
A;Molecule type: protein
A;Residues: 25-41749-760 <HAM>
R;Experimental source: migration-stimulating factor purified from medium conditioned by
R;Sato, T.; Hong, M.H.; Jin, C.H.; Ishimi, Y.; Udagawa, N.; Shinki, T.; Abe, E.; Suda, T
FEBS Lett. 285, 21-24, 1991
A;Title: The specific production of the third component of complement by osteoblastic ce
A;Reference number: S16189; MUID:91293304; PMID:2065778
A;Accession: S16189
A;Molecule type: protein
A;Residues: 25-31 <SAT>
A;Accession: S16189
A;Status: Preliminary
A;Molecule type: protein
A;Residues: 671-677, X', 679-680 <SA2>
R;Fey, G.; Domdey, H.; Wiebauer, K.; Whitehead, A.S.; Odink, K.
Springer Semin. Immunopathol. 6, 119-147, 1993
A;Title: Structure and expression of the C3 gene.
A;Reference number: I49563; MUID:84045280; PMID:6356427
A;Accession: I49563
A;Status: Preliminary
A;Molecule type: mRNA
A;Residues: 25-136, 'X', 138-240 <FEY>
A;Cross-references: GB:M35659; NID:G192280; PIDN:AAA37339.1; PID:G192281
R;Fey, G.H.; Wiebauer, K.; Domdey, H.
Ann. N. Y. Acad. Sci. 421, 307-312, 1983
A;Title: Amino acid sequences of mouse complement C3 derived from nucleotide sequences c
A;Reference number: I49576; MUID:84201365; PMID:6609661
A;Accession: I49576
A;Status: Preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 658-761 <RES>
A;Cross-references: GB:M33032; NID:G192391; PIDN:AAA37378.1; PID:G192392
A;Comment: Complement C3 contains two chains, formed by removal of four residues and lin
alternative complement pathways releases the C3a anaphylatoxin from the amino end of b
native-complement-pathway C3/C5 convertase.
C;Comment: C3a anaphylatoxin is a vasoactive peptide and a mediator of inflammation.
C;Comment: C3b, with its highly reactive thiol group, binds to the surface of foreign pa
e classical-complement-pathway C3/C5 convertase. The activity of C3b is regulated by p
C;Comment: The major site of synthesis of this plasma protein is the liver.
C;Genetics:
A;Introns: 27/2; 90/3
A;Note: the list of introns may be incomplete
C;Superfamily: alpha-2-macroglobulin
C;Keywords: acute phase; complement alternate pathway; complement pathway; glycoprotein;
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-666/Product: complement C3 and C3b beta chain #status predicted <C3BB>
F;25-666/749-1663/Product: complement C3 #status predicted <CC3>
F;25-666/749-1663/Product: C3b #status predicted <C3B>
F;571-1663/Product: complement C3 alpha chain #status predicted <CC3A>
F;571-748/Product: C3a anaphylatoxin #status predicted <C37>
F;749-1663/Product: C3b alpha chain #status predicted <C3BA>
F;946-1303/Product: C3dk fragment #status predicted <CDK>

F;1002-1303/Product: C3d fragment #status predicted <CD3>
F;1424-1457/Region: properdin binding
F;559-816,626-661,693-720,694-727,707-728,873-1513,1101-1158,1358-1489,1389-1458,1506-15
F;748-749/Cleavage site: Arg-Ser (C3 convertase) #status predicted
F;939,1617/Binding site: carbohydrate (Aen) #status predicted
F;1010-1013/Cross-link: thiolester (Cys-Gln) #status predicted
F;1303-1304/Cleavage site: Arg-Ser (complement factor I) #status predicted
F;1320-1321/Cleavage site: Arg-Ser (complement factor I) #status predicted

Query Match	78.43	Score 6749	DB 1	Length 1663
Best Local Similarity	77.13	Pred. No. 0		
Matches 1284	Conservative 179	Mismatches 197	Indels	Gaps 5
QY	1	MGPTSGPSL--LLLLLTHLPLALGSPMYSIITPNIRLESEETVLEAHDAQSDVPVTVT	58	
Db	1	MGPASGSQLVLLVLLASSPLALGIPMYSIITPNVLRSEETVLEAHDAQSDIPVTVT	60	
QY	59	VHDPFGKLVLSSEKTVLTPATNMGNVTTPANREFKSEKGNKEVTVQATFGTVVE	118	
Db	61	VQDFL-KQVLTSEKTVLTQASCHLRVSIKIPASKEFNDSKGEHKYTVVAFGETWE	119	
QY	119	KVLVSLQSGYLFIQTDTIVTPGSTVLYXRIFTVNHKLLPVGRTVMVNIENPFGIPVKQD	178	
Db	120	KAMVSPGSGYLFIQTDTIVTPGSTVLYXRIFTVNDNLLPVGKTVVILITPDGIPVKRD	179	
QY	179	SLSQNLQGLVPLSWDIPELVNMQWKIRAYENSPOQVSTSTFEVKEYVLPSEFVIVEP	238	
Db	180	ILSSNNQHGILPLSWNIPELVNMQWKIRAFYHAPQIFSAEFVKEYVLPSEFVRVEP	239	
QY	239	TEKFYVLYNEKGLVITARELYGKVEGTAFVIFGIDGEGORISLPSKRIPIEDGSG	298	
Db	240	TETFYIIDDPNGLEVSIIAKELYGNVDGTAFAVFGVDGDKKLSLAHSITRVVIEDGVG	299	
QY	299	EYVLSRKVLDDGVQNPRAEDLVGKSLVYSATFVILHSGSDMVQAESRGIPITVSPYQHFT	358	
Db	300	DAVLTARKVLMGVPRSNADALVGKSLVYSVTVILHSGSDMVQAESRGIPITVSPYQHFT	359	
QY	359	KTPKFKFGMPDLNVFTNPDGSPARVPVAVOGEVTVQSLTQDGVAKLSINTHPSQK	418	
Db	360	KTPKFKFPANPDLNVFTNPDGSPASKVLVVTQGSN-AKALTQDDGVAKLSINTHPSRQ	418	
QY	419	PLSITVTRTKQBLSEAOATRTMOALPYSTVGNNSNVLHLSVLRTLPGETLNVNELLR	478	
Db	419	PLTIVTRTKQBLSEAOATRTMOALPYSTVGNNSNVLHLSVLRTLPGETLNVNELLR	478	
QY	479	MDRAHEAKIRYTYLIMNKGKLLKAGQVREPQGLVPLSITDTIPSELYAYVTLI	538	
Db	479	TPDGEAKIRYTYLIMNKGKLLKAGQVREPQGLVPLSITDTIPSELYAYVTLI	538	
QY	539	GASGQREVVDVSVVDVDSKCVGSLVVKSGQSDRQPVPGQMTLKIEGDHGAERVLAV	598	
Db	539	GASGQREVVDVSVVDVDSKCVGSLVVKSGQSDRQPVPGQMTLKIEGDHGAERVLAV	597	
QY	599	DKGVFVLNKNKLTOSKIWDVVEKADICTPGSGKDYAGVPSDAGLFTSSGGQTAQRA	658	
Db	598	DKGVFVLNKNKLTOSKIWDVVEKADICTPGSGKDYAGVPSDAGLFTSSGGQTAQRA	657	
QY	659	ELQCCQPAARRRSVQLTEKMDKVGKY-PKELRCCEDGMRENPMRPFSCORRTFISLG	717	
Db	658	DIETKPAARRRSVQLTEKMDKAGQVTDKGLRKCCEGDGWRDIFPMYSCORRAILITQ	717	
QY	718	EACKKVFDDCCNYITELRRQHARASHGLARSLNDDEDIABENIVSRSEFFESMLWNVED	777	
Db	718	ENCIRAFIDCCNHIKTLREQHRRDHVGLARSELEEDIIPEDIISRSHFFQSWLITEE	777	
QY	778	LKEPPKNGISTKLNIIFLKDSITTWELLAVMSDKGICVADPEVTVMQOFFFDLRLPY	837	
Db	778	LKEPEKNGISTKMNIFLKDSITTWELLAVMSDKGICVADPEVTVMQOFFFDLRLPY	837	
QY	838	SVVRNEQVEIRAVLYNQRQELKVRVELLHNPAPFCSLATTRKHQQTITIPPKSSLVP	897	
Db	838	SVVRNEQVEIRAVLYNQRQELKVRVELLHNPAPFCSLATTRKHQQTITIPPKSSVAVP	897	

1142 ALTAFLVLSIQBAKOTICEQVNSLPGSITKAGDFLEANYMLOBSYTVTAIAGYALAOGR 1201
 1125 YLFAFLVALLBSKTCNDVNSLSSKATNYLLKYEKLQRPYTTALTAYALAAQ 1184
 1202 LKGLPLNKLUTAKDNRRNEDPKQLYNVEATS YALLALLQLQDFDVPVVRWLNQRY 1261
 1185 LNDRLVMAASTGRD--HWEYNAHNTNIEGTSYALLALLKVKKFDQGPVVRWLQNF 1242
 1262 YGGYSGTATFVFOALAOYKQADPHQELNLDVSLQIPSSKITHRIHWSASLRS 1321
 1243 YGTYGQTATVAFQALAEYELQWPTHKDLNLDITIELPREVPIRYINENALLART 1302
 1322 EETKENEGFTVTAEGQGTLSVVTMYHAKADQ--TCNKFDLVKVTIKPAPETEKRPQDA 1380
 1303 VETKLNQDITVTAASGDKATWILTFFYNAQLQEKANVCNKPHLNVSVB--NIHLNAMGA 1359
 1381 KNTWILEICTRYRGDODATMSITLDSMTGFPAPDDDLKOLANGVDRIYSKYELDKAFSD 1440
 1360 KGALMLIKICTRYLGEVDSWTIIDISMLTGFLPDAEDLTRUSKGVDRYISRYEYDNNNAQ 1419
 1441 RNTLIYDKVSHSEDCCLAFKHQYFNVFELIQGAVKYAYAYNLESCTRFYHPEKEDG 1500
 1420 KVAVILVKNVSHSEDECLHFKLKHFEVGIOPGSKVSYVYNLDEKCTKFEHPDKGTG 1479
 1501 KLNKCRDELCAEENCIFQSDDKVLEBRDLKACPGVDYVYKTLKVKQLNSDFDE 1560
 1480 LLMKICIGNVCRGAGTCSLHQBRIIDVPLQIEKACETNVDYVYKTLKRIEEOGNDI 1539
 1561 YIMAIETIKSGSDEYVQVQRTFISPIKREALKLEKHYLWGLSSDFWGEKPNLSY 1620
 1540 YVMDVLEVIKQGTENPRAKTHOY--SQRKQCALNKLKNDYLLIWSRSDDLPTKDKISY 1599
 1621 IIGKDTWVHEPEDECODEENKQCOQDLGAFTESKVVVPGCP 1662
 1600 IITKNTWIERPHEDECOEEFQKLCDDPAQPSYTLTFEGCP 1641
 RESULT 8
 151339
 complement component C3 - rainbow trout (fragment)
 C/Species: Oncorhynchus mykiss (rainbow trout)
 C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
 C/Accession: I51339
 R/Lambert, J.D.; Lao, Z.; Pang, J.; Alsenz, J.
 J. Immunol. 151, 6123-6134, 1993
 A/Title: Third component of trout complement. cDNA cloning and conservation of functional
 A/Reference number: I51339; MUID:94065166; PMID:8245455
 A/Accession: I51339
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-1620 <LAW>
 A/Cross-references: UNIPROT:P98093; GB:L24433; NID:9431607; PIDN:AAB05029.1; PID:9431608
 C/Suprafamily: alpha-2-macroglobulin
 Query Match 40.1%; Score 3452.5; DB 2; Length 1620;
 Best Local Similarity 44.0%; Pred. No. 6.1e-203;
 Matches 728; Conservative 326; Mismatches 528; Indels 73; Gaps 34;
 44 LEAHD-AQGDVPVTVVHDFPDKLVLSSEKTVLTPATNMGNTVTFIPANREFKSEGR 102
 2 VESQDHVGGLPLNKKVIMKVKHPTQSKELASKSVLVDQANNFQMTQLVIQRGLVDDP-Q 60
 103 NKEFTVQATFGTQVBEKVLVLSQGYLFIQDRTIYTPGSGTVLRYFTVNHKLLPVGR- 161
 61 KQVVLQAQPPDELLEKVLVLSFGSYFIQDRTIYTPASTVHYRVFSMTPLGLEPTRE 120
 162 -----TVVNTENEGIPVKQDSLUSSQNLGVLPDWDPELVNMQWKIRA 208
 121 IFEDQEVAKNKEIAVGEVIMTENITIFREIVNPKD--GVKSGQFKLFDIVSFGTHVVT 178
 209 YVENSQQVFTSEFEVKEYVLPSPFVIVEPTEKFFIYVINEKGLEVITITARELYGKVEGT 268
 179 RFQSTPQKTFSEFEVKEYVLPSPFVSLTPAKAFFYV--DDNDLTVDITARYLYGKEVTGT 237
 269 AFVITGTDGP-QRISLPESLKRIPEDSGSEVVLARKVLLDGVQNPRAEDLVGKSLYVS 327
 238 GYVFGVITTESEKKSFPASLQRFVEIKDGGVACLKEHITQT--PKHDLVKQSFYVS 295
 328 ATVILHSGSDVQAESGIPVITSPYQIHFTKTPKPKGMPDPLMVFTNPDGSPAYRV 387
 296 VSVLTGEGEMVBAEKGIQIVTSPYSILFKRTPKPKGMPDVSIVITNPDNSPAICV 355
 388 PVAVCGEDTVOSLTQDGVAKLSINTHPSOKPLSIIVTKKQBELSABEATMTQALPYV 447
 356 EYEVTP-PHAKGVTRANGFAKPLNTVASATELVIIVTKPDGPPRQOQTGGTGMALPYR 414
 448 TVGNSNNYHLISVLRTELPEGETLNVNFLRMDEA--HEAKIRYVYTYLIMNKRILKAGR 505
 415 T--STKNPLHVGVDNELKIGDKIDNLGPTTIPNHD-----LTYMFLSRGQLVKVGR 467
 506 QVRPQGLVLPISITTDIPSPRLVAYVTLTGASQREVVADSVVWVDKSCVGSUVV 565
 468 FKRG--GNALVTLSPVSKELLPSFRIVAYYHV---GAADLVADSVVWVDIKVSCMGSLKV 522
 566 KSGQSEDRQPVPGQOMTLKIEGHRGVVLVAVDKGVFVINKKIKLTKISKIWDVVEKADI 625
 523 TSTRPKASYE--PRAPSLTITGDPGAKVGLVAVDKGVVILNSKHLRTQTKIWDTEKHT 581
 626 CQTPGSKDYAGVSDAGLTFTSSGQQTACRAELQCPQAPARRRSVOLTETKMDKVGK 685
 582 CQTAGGAGNMGVFDYAGLVFETNAGTGIRTDPSCP--VSSRRRAVTVISVITSMASK 640
 686 YPKELRKCCEDGRENPMRFSCQRTTFISLGBACKKVFLDCCNYITEL--RRQHARASH 743
 641 YHGLAKECCVGDGMDTNGYTCORRAQYISDGVCVQAFLVCC--TEMASKKIESKQDA 697
 744 LGLARSLNDEB---IAENIVSRSEPPESWLNWVDELKEPP---KNGISTKLM-NIFLK 796
 698 LLLRSSEDDDDDAYMSESDIVSRQSPESMMWEDNLPECPAQNKHCSTSVIRNNFLK 757
 797 DSITITWEILAVMSGDKKIGICVADPPFVTVMDQFFIDRLPYSVVRNEQVEIRAVLYNQ 856
 758 DSITITWQITA-SLSKTHGICVADPPFEMVLKEFFIDLKLPSAVRNEQLEVKAILHNYSE 817
 857 NOELKVRVLLHNPAPCSLATTKRHOOTITIPKSSLSVPYVIVPLKGLQEVVEKAAV 916
 818 D-PIIVRVELMENGECVSSASKGKRGQEVNMPDMSTRVVPVVIIPMKLGLSHSIEVKASV 876
 917 YHFTISDGVKSLKVPVPEGIRMKNTVAVRTLDPERLREGVQKEDIPPADLSDQVPTES 976
 877 KNSGSDGVKRDVAVAEGLVKKETNV--LLNPVKHG--GEQTSHPGSGVPRNQVNSDA 933
 977 ETRI-LLOGTPVAQMTEDAVDAERLKLIVTPSGCGEQNMIGTPTPIAVHVLDETEQWE 1035
 934 DTLISVTAGEQTSVLVSQLSGSLGLIVQPGCGEQNMIMYTLPIVIAHVLNDTKKWE 993
 1036 KFGLEKQCALLELIKGYTQQLAFRQPSAFAAFVFRAPSTWLTAYVVKVFSVLAVNLIAI 1095
 994 DIGLDRNTAKYINIGVQRLAYRKEDSGYAAVWSRQSTWLTAYVVKVFSVAMSSTLIS 1053
 1096 DSQVLCGAKWLIILEKQKPGQVQEDAPVHQBEMIGGLRNNNEKVALTAFVLISLOEAK 1155
 1054 QENVLCTAVKWLIIINTQPDGIFNEFAPVIAHEWTGNVR--GSDNDASMTAFVLIAMQES 1112
 1156 DICEQVNSLPGSITKAGDFLEANYMLOBSYTVTAIAGYALAOGRKGLPLNKLUTTAK 1215
 1113 SVCQSVNSLPGSNAKAVAVLEKELPHLTPNPAVAMTSYALANAGKLNKETLLKSPQL 1172
 1216 DKRWEDPKQLYNVEATSYALLALLQLQDFDVPVVRWLNQRYVGGYSGTATMY 1275
 1173 D--HWPVGGVQYVTEATSYALLALLKVKAFESAGPIVRLNKKQKVGCGYSGTATMY 1230
 1276 FOALAQYOKDAPDQELNLDVSLQIPSSKITHRIHWSASLRSSEETKE-NEGFTVTA 1334
 1231 FOAVAEYWSHVKDLKDFDLNINLEAVAGRASVTKWSNNKQPHTRDKVNSIDKDLTVKA 1290

Db 230 PTEEVINPQRFVHI-NDEEFWDITANFYFNQELVSGTAVRYFLENGD---VPLVD 284
Qy 289 -KRIPIEDSGGEVVLSPKVLJOGVQNPRAEDLAVGSLVYSATVILHSGSDMWQAEBSGP 347
Db 285 SSSTTVAGEGSLKKEKLLK--LFFNAKOLLAFSLTKITVILSSQAETEEAELVGIK 342
Qy 348 IVTSPQIHTFKTPKFKGMPFDMVFNFTPDGSPAYRVPVAVQ---GSDTV---OSLT 401
Db 343 IVERSQITATKTSRYFKEPELFFIQVEVSNADGSPSKVEDVWAKVQVGSATINPQKMT 402
Qy 402 QDGVAKLSINTHPSKPSLITVRKQELSEAGATMCAALPYSTVGSNNYVHLISVL 461
Db 403 DSNGLTSFTV-TPPNVQUTVVRIDERHPSN-EGELVYTAQKVA---SASTMHDVT 456
Qy 462 RTELPGETLNVNFLMRDRAHEAKIRYVYILMNKRLKAGRVPRPGDVLVPLSL 521
Db 457 RI-MRLGETLNVFLTAKTQLN--AVTHFTYMLTRGVIVKTRNKTESGGGPNRIP 513
Qy 522 TTDPIFSEFLVAVYTLIGASGOREVWVADSVWVDKSCVG--SLVVKSGOSEDQPVPGQ 579
Db 514 TPDMAFRFLAYIPLGG---EIVADSVTVETELCKSQVSLKGRPTLE---PKA 565
Qy 580 QMTLKIEGHDGARVLVAVDKGVFLVANKKNTQSKINDVVEKADIGCTPGSGKDYAGVF 639
Db 566 MLTDLIGSPDARVGLAVDQAVYAVNRKHLTQDRVWKAMETFTDGTABGGAGRPQVF 625
Qy 640 SDAGLTFTSSGGQTAQRAELQCPQAPARRRSVOLTEKMDKVCKYKPELKRKCCEDMR 699
Db 626 SDAGLALITSGKLTNDRSEIGCPKVPKPKQLSMLQIRRE-AEKTOEPRKCCVDGLK 684
Qy 700 ENPMRFSQORTRFISLGBACKKVLDCNYYITELRQARASHLGLARSNLDE-DIAE 758
Db 685 MSPTQGCCERLKRVTGPKRECVDAFLQCKAAEYRKESLCAKTVLRRNDFMELDLNNE 744
Qy 759 ENTVSSEPPESLWNVDELKPPKNGISTKLMTFLKDSITTWELLAVNSDKGKICVA 818
Db 745 DEVMNMYFPQSGWYN--KYKNSCKYGRHPQI-RLQLPDTITWNNQVSLSKTRGVCLA 801
Qy 819 DPEVTMDDFFIDLPLPSVVRNEQVEIRAVLYNQRQBELKVRVELLNHPAFCSUAT 878
Db 802 DPLLVSRTKDFKILPLPSVVRNEQVEIRAVLYNQRQBELKVRVELLNHPAFCSUAT 878
Qy 879 KRHQOTITIPKSSLSVPVIVPELKTGLQEVVEKAAVYTHFISDGVKRSKVVPEGIR- 937
Db 861 GAPSQKSVKGAUVSPFVPLKIGEHHSIRSVYGRFTGDCGVQKILRVAPEGVRD 920
Qy 938 -MKNVAVRTLDPERLREGVOKEDIPPAADLSQVDPDTESETRILLQOTPVQWTEDAVD 996
Db 921 IRSERSVHEBERETP---FINETISP---DVVPNSDVLTFISVKGDELAETWVNCGL 972
Qy 997 AERLKLHVLTPSCGQBNMIGMTPTVIAVHYLDTEQWEKPGLEKROGALELIKGYTQ 1056
Db 973 AKSINLQIPTGCGEONMKNAPTTLILVLDVQWEKGLHRREAEITGLAQGSRE 1032
Qy 1057 LAPRQSSAPAPKAPSTWLTAYVVKVFSLVANVLIAIDSVLCGAVKMLILEKQKPDG 1116
Db 1033 LSYRKADHSYAAPIKRPSTWLTAFVVKVYSLAKREVIIVDQELCGPVEWIKNRQNSDG 1092
Qy 1117 VFQEDAPVHQEMIGLNRNNEKDMALTAFLVLSQEAKDICEQVNSLPGSITKAGDGL 1176
Db 1093 SYREDGPVIREMCGGV-GGTEGHVSMATFILIIGQQAQEGYCVSVPNYKSNMRAVQFL 1151
Qy 1177 EANYMNLQSYTVAINAGYALAQMGRLKGP-----LLNKLFTTAKDKNRW---EDP 1223
Db 1152 ASKVSOLKMYTIAITRYALA---LQDPESEAAHSWKKLENKTRTFESKGRHYWKAET 1207
Qy 1224 GKOL---YVNEATSALLALLQKDFDPVPVFWNLNEQRYGGYGSSTOATPMWFOAL 1279
Db 1208 SHVLRMSAISVEATAGLLTYLRKKDYESAREIVDMLTQRYGGYGSSTOATPMWFOAL 1267
Qy 1280 AQYQKADPHQELNLDVSLQFSRSKITHRLHWESASLLSBETKENEGFTVTAEGKG 1339

Db 1268 AQYKMDSSSKELIDVQLLEITSPPKONPEKMKIITEETFVQEPHKIPPGGNITIKASGRGT 1327
Qy 1340 GTLSVWVWYHAKAKDQJLTCKNFKDLKVITKAPETKRPQ-----DAKN----- 1382
Db 1328 FTLSIMSVFNKVAAPSSKSCSTFDLKVMTTEADGES-PQRLGWFDGKRRRRDIDEGG 1386
Qy 1383 ---TWLEICTRYRG-----DQATWISILDISMWTGFAPTDDDLQJANGVDYRYISKTEL 1434
Db 1387 VEAVYRMMCTRYKPRKEDLSSESGMTIIEVNMLTGFIIDKNDLIQKESVDKYISNYEI 1446
Qy 1435 DKAFSDANTLIYLDKVKSHSEDDCLAFKVHOYENVELIOPGAVKVYAYVNLSESCTRFVH 1494
Db 1447 TD-----SVLLIYWDKVPSTEDYCFAPKIKQMLRSDMIQFVTVASYDYVSPADKCTRLYN 1501
Qy 1495 PEKEDGKLNKCRDELRCRAEENCFTQKS--DKVTLEERLDKACBPGVDYVYKTRLVKV 1552
Db 1502 LPGAIVYVSLPCLQNDLCCQVEVSPKPKFDTSIITLHROEAAACVAGIDYAYGIVDNR 1561
Qy 1553 QLSNDFDEYIMAEQITIKSGDE--VQVGOQORTFISPIKCREALKLEKPKHYLMWGLSSDF 1611
Db 1562 TEVGSFVYVYVNIQTVIKSGDQCAIQPKATRLFIVTRS CDGLGHEVTPRQYLLMGRKGET 1621
Qy 1612 WGEKPNLSYIIGKDTWVHWPEDBODEENQKQODLGAFTESMVVFFC 1661
Db 1622 KDRNDRFYVLDASSWYEQWPDVEKCNQPNVQTFCAIKREYEFMSQIQGC 1671
RESULT 11
T30517
Complement C3-Q2 - common carp
C:Species: Cyprinus carpio (common carp)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C:Accession: T30517
R:Nakao, M.; Obo, R.; Mutsuro, J.; Fujiki, K.; Nonaka, M.; Yano, T.
submitted to the EMBL Data Library, July 1998
A:Description: Sequence diversity of the third complement component (C3) from a bony fish
A:Reference number: 220840
A:Accession: T30517
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1230 <NAK>
A:Cross-references: UNIPROT:Q9YIA5; EMBL:AB016215; NID:d1253045; PID:d1037610; PIDN:BAAS
Query Match 28.3%; Score 2438.5; DB 2; Length 1230;
Best Local Similarity 40.5%; Pred No 5.7e-141;
Matches 522; Conservative 231; Mismatches 406; Indels 129; Gaps 19;
Qy 388 QSLTQGGGVAKLSINTHPSQ-----KPLSIT-----VTKKQELSEAEQATR 439
Db 446 QDYSREDLVNKLIVYKHPKMDILKSVTLTAANGFQILTDFKATKDPEDLDEQQAVR 105
Qy 440 TMOALPYSTVNSNNYLHSLVRLTELPGETLNVNLLMRDRAHEAKIRYVYILMNKGR 499
Db 106 KMTAQAVETKGGSSQNYLHIGIDAELQIGDSVRV--YLTNGQIPGVKQDQFTYMLSKQG 163
Qy 500 LLKAGRVREPGDQLVPLSITTDIPSPRLVAYYTLIGASGQREVVADSVWVDKSC 559
Db 164 IVKVDKFRKR-GQSLVTLPTVTVKDMPSPFRFVAYTHV----GSSEWSDSVWVDKDT 218
Qy 560 VGLSWKSGQSEDRQPVPGQMTLKEGHDGARVWLVAVDKGVFLNKKKNTQSKIWVY 619
Db 219 MGTQIKVKDKLNIYDT-GDEVKLQITGDPGKVLVWVDKAV---NKNRLTQIQIMDV 273
Qy 620 VEKADIGCTPGSKDYAGVPSDAGLFTTSSGGQTAQRAELQCPQAPARRRSVQLTEKR 679
Db 274 IEKHIDIGTAGGGRDSMGVFTDAGLAFESNAGTGNTRTTPDCPFLSKERRS----- 326
Qy 680 MDKGVKYPKELRKCCEGDMENPWFSCQRRTRFISLGBACKKVFJLDCNCCNYITELRQHA 739
Db 327 -----DDDDYVD----- 334
Qy 740 RASHLGLARSNLDEDTAENIVSRSEFPESWLNWYVDLKEPPKNGISTKLWNILFKDSI 799

Db 335 -----SEIVSHQTQFRESNWEIEDLDCDKCSTPATEKV--IYLKDSI 374
Qy 900 TTWEILAVMSDKKGCIVADPFVETVMQDFFIDRLPLFYSVVRNEQVEIRAVLYNQRQCE 859
Db 375 TTWQILAVSLPTLIGICVASEPBMVFKILFIDLKMPYSAVRGEQLEIKAIHNYTNKQ 434
Qy 860 LKVEVELLHPAPCSLATTERRHOQITIPPKSSLSVPYVIVPLKGTCEVEVKAAVYHH 919
Db 435 -KXVVEPWEIEDVCSFKAGKRTTVSVKUSSISVSIVLIPMLGNHMEVKSAYDA 493
Qy 920 FIDGVRKSLKVVPEGIRMNKTVAVRTLQPERLREGVQKEDIPFADLSQDQVDPTESTR 979
Db 494 IYTDGVRKQLKVVSEGLTVSHKGNVELNPVKNQDPLVFKSVIDA---DRLEDPADMY 550
Qy 980 ILLOGTPTVAQWTDVADAEELKELIYTBSCGQONMIGTPTVIAVHYLDTEQWEKGL 1039
Db 551 IAITGEIITQVEOASGSMGLIYQPSGSGQONMMLTFLIATHYLDSTSQWDTVM 610
Qy 1040 EKGQALELILKGYTQQLAFRQSSAPAFVFRAPSTWLTAYVVKVFLAVNLIAIDSOV 1099
Db 611 ERNEAVNYINTGYRQLGYRKSDGSGYAAWTHRPSTWLTAYVAKVFSMANNIVNEV 670
Qy 1100 LCGAVKLLILEKQKQDQVCEADPVTHQEMIGGLRNNEKDMALTAFLVLSIORAKDICE 1159
Db 671 ICSALKKLLIHKQTLGSKFEDSAVQGENWGIQ--SKDASLTAFAVIAQEAELICA 729
Qy 1160 EQVNSLPGSTKAGDFLEANYMNLORSYTVIAGYALAQWGLKGLPLTKPITTAOKNR 1219
Db 730 GSVASLHESIRKAVSFLEGLPLQLTNPYAVAMMSYAMANEKLNKILMKHSSQEAQRS 789
Qy 1220 WEDPGKOLYNVEATSYALLALLOLQDPPVPPVVRWLNQRYGGYGSQATQATMVFOAL 1279
Db 790 WTVPGQHHLEATAVALALVAKAFDRAGEAVHWNQOQHYGSGTQTATLVFOAV 849
Qy 1280 AQYQKADPADQOELNVLQPSRSKITHRIHWEASALLRESEETKENEGETVTAEGKQ 1339
Db 850 AEYRQVQDQNFENLDELVSAGRGKPVRYTIKRENAHLTWGKMEINKEFNVTARGTGR 909
Qy 1340 GTLSVVMYHAKADQLT--CNKFDLKVITKAPETEKRPQDAKNTWI---LEICTRYEG 1394
Db 910 ATLSVLILYARPEKKSQDCTFDLTKM-----EKNEAKHGHTIASYKLTMDPIYKS 962
Qy 1395 DQ-DATMSILDI SMGTGAPDPTDQLKQANGVDRIYSKYELDKAFSDRNTLIYLDKYSH 1453
Db 963 DKTATMTILDVGLPTGFEVESDLKQLSGSKERYIQKPEMNKVLSEKSLIYLDKYSH 1022
Qy 1454 SEDDLAFKQVQVENVELIQGAVKYVAVYNLEESCTRYVHPEKSDGKLNKLCRDELQRC 1513
Db 1023 KEKVVISFRHQJYVUGLLQPAATVIEYISPDARTMYPHERDAIYRLUCKGDLCLC 1082
Qy 1514 AEENCFIQSDDKVTLEERLDKACEFGVDYVYKTRLVKVLQNSDFDEYIMAEIQTKSGS 1573
Db 1083 AEENCSFQKNN--VRDEERLNRVCEPQMDYIKVTVVGMDLKQDSDIYDMKVELVKEGT 1141
Qy 1574 DEVQVGQORTFISPIKREALKLEEKHYLMWGLSSDFWGEKENLSYIIGKDTWVHWPE 1633
Db 1142 DEDAGKVRQFLAPSCREHLGLVECKSYLNGRSDDLPELQSGSLQYVFGQWVEYIYPT 1201
Qy 1634 EDEQDEENQKQCODLGAFTESMVVFGC 1661
Db 1202 REESQTRHRRVIGISELQNSL-RYGC 1228

RESULT 12
S21045
complement protein homolog - inshore hagfish (fragment)
C:Species: Eptatretus burgeri (inshore hagfish)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S21045
R: Ishiguro, H.; Kobayashi, K.; Suzuki, M.; Titani, K.; Tomonaga, S.; Kurosawa, Y.
EMBO J. 11, 829-837, 1992
A: Title: Isolation of a hagfish gene that encodes a complement component.
A: Reference number: S21045; MUID: S2192016; PMID: 1372251

A: Accession: S21045
A: Molecule type: mRNA
A: Residues: 1-1620 <ISH>
A: Cross-references: UNIPROT: P98094; EMBL: Z11595; NID: G62774; PID: G62777
C: Superfamily: alpha-2-macroglobulin
C: Keywords: plasma

Query Match 25.0%; Score 2148; DB 2; Length 1620;

Best Local Similarity 31.7%; Pred. No. 5.9e-123; Indels 124; Gaps 44;
Matches 535; Conservative 338; Mismatches 60

Qy 13 LATHPLALGSPWYSIIPTNIRLESSEETMVLAEADQGDVPTVTVHDFPKKLVLSS 72
Db 1 VLVIAFAATSS--YDDLAVAILMVDQKK--ITEVH-----VLVNPHTGATL--DE 45
Qy 73 KTVLTPTATHMGNVFTTPANREFSEKGRNKFVTVQATF--CTQVWEKVLVLSQSGYLF 131
Db 46 KKVKLQWDMKF--IAFTKLQVTPKEVEKQEDFVLMVKWDGQGHMELIDPLTSRRGLV 103
Qy 132 IQTDKTIYTPGSTVLYRIFTVNHKLLPVGRITVMNIENPEGIPV---KQDSSLSSQNLGV 188
Db 104 AQTDQFIYTPNDVNIRLFPVTRQLNPLLSLVVDIMNPDGVVVDRIEKNAFEVKEWML 163
Qy 189 LPLSWDIPELVMGQWKIRAYYENSPOQVFSTETEVKEXVLPSEVIVEPTEKYYIYNE 248
Db 164 RP--PHVPAITSLGDMKIVSMKDKPQFNYSYSGFKVEEYVLPFDVSTISQPYLHYV-D 220
Qy 249 KGLEVTITARFLYKKEVGTAFVFIQDGEQRISLPESLKRIPIDEGSGEVLRSKVL 308
Db 221 KAPTTHIKAMHYGKPVMGRAVRYGVKHQSKRTLLSTSSALARPEQGEAMHTLRQKHL 280
Qy 309 DGVCNPRADLQKSLYSATVILHSGSDMVQABERSGPIVTSYQIHFTKTPKYEFGM 368
Db 281 EQVPDPLK--LIGQSUYVEASVISSDAEIEISILDDIPIVASPYISKXWTVFFPKGV 338
Qy 369 PEDLAVFVTPDGSFAYRVPVAVQ-----GEDTVQSLTQGDGVAKLSINTHSORPLS 421
Db 339 FYIKVVLNLPDGSFASGVPVVKVSPFSSGSGNWTQKTKTMDNGIAMQTINTARNKKLN 398
Qy 422 ITVTRKQBELSBAEATRTMQALPYSTVGNNNYHLVSLTELEPGETLVNVLNLLMDR 481
Db 399 IKVQTEDELEQSQAEASFTIASYSS--PSGFIHLNAHREVKSPEG--HIVFDVFIKS 454
Qy 482 AHEAKIRYTYILMKNRGL---LKAQRQVRPQDQVLVPLSLTITDFIPSPRLVAYITLI 538
Db 455 AAKDHVLFHNYLMIKNGIHNFLQSGRK----GDTTSVSLLLTPELVQFELVAFITL- 508
Qy 539 GASQREVVDVSWVDVVDKDS CVG--SLVYKSGQSEDRQPVFGQMTLKIEGDHGRVTVV 596
Db 509 -PSG--ELVADSIIDVXDSCAKLSLDVAGK---RLFSPRDNVNFDSGESDSWAVG 562
Qy 597 AVDKGVFLNKKNLTKQSKINDVVEKADIGCTPGSGKDYGAVFSDAGLTFTSS--SGQQT 654
Db 563 VVDXAAVYLDKXKLTANKVYKAMEASDLGCSVSGSGTKPLVFRDAGLAIWAKISGMD 622
Qy 655 AQBAELQCPQAPARRRSVQLTEKMDKVYKPELRCCEDGMREPMRSCORRTFI 714
Db 623 VK--DPCGNHTRRRRLVL--EIAIERKASTYPAELRKCRCDAIAESPLRUSCEERTKI 679
Qy 715 -SLGEACKKVFLLDCNNYITELRQARASHILGLAFSLNLDIEDIIAEENIVSRSEFPGLW 773
Db 680 HDEGECQSTFLECCCKHVEEELLIAVEBEDDLGRSQGEDFMJQESQVIRSHFESFMV 739
Qy 774 NVEDLKEPPNGIS--TKLMNIFLKSITTWELLAVSMDDKGLCVADPPFVTVMQDFFI 831
Db 740 EIKUSRAENKSRITKKN-----PDSITTWIDQAVEVSQSGKLCVGSLSLTTFVKQFFL 795
Qy 832 DLRLPYSVVRNEQVEIRAVLYNQRQELKVRVLEHNPAPFCSLATTKRHQQTITIPP 891
Db 796 KVHTPYALKQYEQVELRVYVINY--NNQVKGIEIQKCGDGICTDAEQNEPLKSRFAVEKN 854
Qy 892 SSSLSEPVYVPLKTLQGEVEVKAAVYHHFISDGVKSLKVVPEG--IRMNKTVAVRTLD 949
Db 892 SSSLSEPVYVPLKTLQGEVEVKAAVYHHFISDGVKSLKVVPEG--IRMNKTVAVRTLD 949

Db 855 SATSPFVWVPLSSSDSSVSVLAVFGSDVDHDAVEKDLRVMPGEGNEMERSHSVQ---P 911
 Qy 950 ERLGREGVQKEDIPADLSDQVDPDTESETRILLQTPVAQMTEDAVDAERLKLHIVTPSG 1009
 Db 912 RRRG--GQGVIVVDNETPQNVVPGT--EMGAFLSAQGNLVAETIQNTLKGSKIINLLRPRG 969
 Qy 1010 CGSQNMIGTPTVIAVHYLDETEOKEG--LEKROGAELEIKKGYTQQLAFQPSASF 1066
 Db 970 CGSQNMVYTSITVWARYLNRSDQNKMGDPOLKRR--SFDITSGFASQUTRKPDYSY 1027
 Qy 1067 AAFVKKRAPTWTAYVYKVFSLAVNLIAIDQVLCGAYKWLILKQKPDGVDQFEDAPVH 1126
 Db 1028 AAMLHRSASTLTAFVAKVFSQARQLVFPVSEICGSVRL--MRQDKDSFLESKEVHV 1086
 Qy 1127 QEMIGLNRNNEKMAITAFVLSIQEAKDICEBOVNSLPGSITKAGDFLEANYMNLQRS 1186
 Db 1087 LNMGGVTK---VVLTSFVFIALLARESCINEVEGTVVVKAHGYLTSQAMGLED 1142
 Qy 1187 YTVAIAGYALAQMGRKGLPLNKFTTTAKDN-----RWDPGKQLYNVYATSVALLA 1239
 Db 1143 FPLAITAYALS-LWKVSDGAAKVTMTLTKTSGLQTEELIHWGSKNGKAAAVESTAYGLLA 1201
 Qy 1240 LIQLKDFDFVPPVWLNQRVYGGYGSTQATFVFOALAQYOKDAPDHCENLNDVSLQ 1299
 Db 1202 AICHESGEAETATNLSOSATFGYFQSTQDTWALQALTGFSCQSRMKMDLSFKIR 1261
 Qy 1300 LP-----SRSSKITHRIHWESALLRSBETKENEGFTVTSAGKQGTLSVVVTHAKAD 1354
 Db 1262 AEENGVDKFEQITN---DNAFVQKPKVPVHVGQITVTSAGTGQGITVFKKYREKVI 1317
 Qy 1355 QLTCKNFKDLVTIKPAPETEKEPODAKNTMILEI---CTRYRGDQDATMSILDISMMTG 1410
 Db 1318 KDCQKFSLEITNLDNQVRRROSINPBNVVRPFCFRYLNRQSPGMVMDISLPTG 1377
 Qy 1411 FAPDPTDLKQLANGVDRIYSKVELDKAFSDRNTLIYLDKVSHEDDCLAFKVHQYNVE 1470
 Db 1378 FEAKKXDLDMKNLVDNYVQVEIRP---GRVFLYLDKVNKDKCNVGFRLNQVPSN 1432
 Qy 1471 LIOPGAKVYVAYNLEESTRYPHPEKEDGKLNKLCRDELCKCAEENCF-IQKSDDKVTL 1529
 Db 1433 LVLPVTATVFEYEPFRCSKYHPPKMEVNP--DASCHGNI CNCLQRHCVELKGMADERN 1491
 Qy 1530 EERLDKACBPGVDVYKTLRVQLSNDFDEYVNAIETIKSGSDE-VQVGQORTFISPI 1588
 Db 1492 ADRNGNACR--AEVFEIIGVTKVTAGYININAAKTLVKKGMDOQAINVGARRSFVPM 1549
 Qy 1589 KCREALKLEKKHLMWGLSSDFWGEKPNLSVYIIGKDTWVGHWPEED-----ECOD 1639
 Db 1550 HCKNLNVPGDIYLVGMHNAHNSORTQVLTSDTWFEKFPLESVCRLPSPASCQV 1609
 Qy 1640 EENQKOC 1646
 Db 1610 SENFKGC 1616

RESULT 13
 T30885
 Complement component C3 homolog - African clawed frog
 C;Species: Xenopus laevis (African clawed frog)
 C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
 C;Accession: T30885
 R;Mo, R.; Kato, Y.; Nonaka, M.; Nakayama, K.; Takahashi, M.
 Immunogenetics 43, 360-369, 1996
 A;Title: Fourth component of Xenopus laevis complement: cDNA cloning and linkage analysis
 A;Reference number: 220919; MUID:96186527; PMID:8605056
 A;Accession: T30885
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-1683 <MOR>
 C;Cross-references: UNIPROT:Q91741; EMBL:D78003; NID:g1183856; PID:BAA11188.1; PID:g118
 C;Superfamily: alpha-2-macroglobulin

Query Match 24.6%; Score 2122; DB 2; Length 1683;

Best Local Similarity 31.58; Pred. No. 2.5e-121;
 Matches 554; Conservative 328; Mismatches 704; Indels 170; Gaps 49;

Qy 1 MGTPSGSLLLLL---THLPALGSPMYSIITPNILRLESETMVLBAHQDQVPVTV 57
 Db 1 MGP---PPLLYALVMAAVCMVAKAQPNFLWAPRIILHVGVKEVICIGIPEWPGQSLPAG 57
 Qy 58 TVHDFPGKLVLS-----SEKTVLIPATNHMGVNTFTIPANREFKSEK----- 100
 Db 58 DI-----KVELSURNQLSMIGSGSEKFLNKRND-----YTILTMLTSQRLITCK 105
 Qy 101 -----GRNRFVTYQA--TFGTQVVEKVLVLSQSGYLFIQTDKTYITPSTVLYIFFTW 153
 Db 106 LNNRRSGRYQLVVKSDVLGKQKWSIPVAYQRYLFIQTDKSIYTPKRETHIRSPFLD 165
 Qy 154 HKLLPVGRITWVAIENPEGIPVKQDSLSQNLGVLPLSWDPELVNMGOWKIRAYVENS 213
 Db 166 HVLPRTBEQVTLVFNACQGFQVRKVKISKD--SVVADNLQIPDIPSTPGVWRISVHYTDA 223
 Qy 214 PQOVSTEFYKHYVLPSPFEVIVETPKFYIYNEKLEVTITARFLYKKGVEGTAFFVF 273
 Db 224 PETNFTAEFEKVPYLPNFEVKIIP--EVPYFLMTKDSFTFRVEARYVYGEHVAGVGNRV 282
 Qy 274 GI--ODGBOIRISPELKRIPIEDSGSVVLSRKLVDGVQNPRAEDLVGKSLVYSATVI 331
 Db 283 CITRQNG--KRYMLAGLEKQTTLDGTTVTIRWADIKELIQODMG--NLLGTSLSIPAAVSV 340
 Qy 332 LHSQSDMVQAEKGIPIVTPSPQIHFHTKTPKYPKQMPFDLMVFTNPDGSPAPVAVPVAV 391
 Db 341 EKASGVLEKEFTSVKFEVSSPYKLDLSKTRYPITGTPAQIIVVEVSHIDGSPAGVSVL 400
 Qy 392 QGEJT--VQSITQDGVAKLSINTHPSOKPLSTVTRTKQELSEAEQATRTMQALPYSTV 449
 Db 401 SKSNAQIFKYKTDNNGVVAEHNVADEKCDIRKANFDSSESESIT---LLPYTS- 455
 Qy 450 GSNNNYLHLSVLTETELPGETLNVNLLRMDRAHEAKIRYTYLLIMNKGRLKAGQVRE 509
 Db 456 -KASSYFLVAVPNQVLDPGSSFKVT--LKAIISNKQVNEKIYYMVLNKGQLSLDSISRT 512
 Qy 510 PGQDLVPLSITTDTPSPRLVAYTYTLIGASQREVWADSVWVDVKQSCVGLSVVKGQ 569
 Db 513 EVNEMLI---TVKESMTPSPFRVIAYYL-----GSEIISNVWVDVADVCESKELHASK 564
 Qy 570 SEDRPVPGQOMTLKIBGDHGARVILVANDKGVFVNKNKNTQSKINDVVEKADIGCTP 629
 Db 565 ---KILAPGALKLDVTEGTATVLSAVDTAVYILNSKNKLTPOKMFKANNAYDLGCSF 621
 Qy 630 GSGKDYAGVFSDAGLTFTSSSGQQTQRAELQCPQFAARRRSVQLTEKRMKVGVKY-PK 688
 Db 622 GGGKDFINVTDFAGLAFVSSAGY--TQINELGC-RVHQKKKALDFOALTQOKAYSITT 678
 Qy 689 ELRKCCEDGMRENPMRFS--QORTFISLGEACKVFLDCQNYITELRQHA--RASHL 744
 Db 679 ELQRCQGHMMLPFGKMSRVCTKRAARVP--DPTCRKAFILDCEYAEENLKQLTEKRTKQ 737
 Qy 745 GLARS---NLDEDI IAEENIIVSRSPFESLWNVSDLEKPPKNGISTIKLMMIFLKDSIT 800
 Db 738 CFGITQNVGSDDEDFADESDIQIRSFPEFWLWRTVKV---YNGLFSA--VYMPDSIT 791
 Qy 801 TWEILAVMSDKKIGICVADFEVTVMDPFIDURLPSYVRNBEQVEIRAVLYNYNQNEL 860
 Db 792 TWEIQAIGMSREKGFCAEPLKVKVFPDPIYLRVPYSVKRFEQMBELRILYNY--NNKDL 850
 Qy 861 KVRVELLHNFAFCSLATTKRHQ--QTITIPKSSLSVPYVIVPLTKTGLQEVKVAAYVHH 919
 Db 851 EVKYWMECAEDICSPGSDSKPLIKVETVGCANSALPFPVWVPIGKSNPVSV-VALGRS 909
 Qy 920 FDSGVKSKLVKVPVEGRIMNKTVAVRTLDPERLGRGVQ--KEDIIPADLSQDVPTESET 978
 Db 910 FVSDGVKAMKIVKEGASVFEESY--IIPADITRISIDFDEBFP---SNMIPGDPRS 964
 Qy 979 RILLQGTVPVQMTEDAVDAERLKLHIVTPSGCGEQNNIGMTPTVIAVHYLDETEOKEGF 1038

Db 965 SIKVTMDSSMNTINNSLGADGISKLRVPYGCABEQTMISTSPGVVALRYLDHTKWNLLS 1024
 QY 1039 LEKQGALEIKKGYTQOLAFROPSPAFARFVGRABSTWLTAVYVGVFSLANVLTAIDSQ 1098
 Db 1025 PDRKDEGLNRQYGLILQFKADGSGYGLWHLRTSTWLTAFVVKVSLCNYIDVWVE 1084
 QY 1099 VLGCAVKNWLEKQKPGVFCEDAPVHQIMIGGLRNNEKDWALFVLLISLOBAKDIC 1158
 Db 1085 DIRLSAQYL-ATMOKDYGAFQKSVIHQDMLGG-STTIDAESVITAYVTVSLYHSLDSL 1142
 QY 1159 EEQVNSLPGSTTKAGDFLEANYNMLORSYTVATAGYALACMGF--LKGPLNKELTITAK 1215
 Db 1143 SEDNVAVKSKIAVDYLRKGLDITIKHPYSALUTYVALTILTSKOSVLKDKAYNKMLMSRAQ 1202
 QY 1216 -DKRWEDPGKQLY-----NVEATSYALLALQLKDFDFVPPVVRWLNQRYYGGY 1266
 Db 1203 GDFPK-----KELYFGPKGTALAVETTSYVLLITLGRNITKAERMYTWLSRQYGGGF 1257
 QY 1267 GSTQATFMVFOALAQY--QKXADPHQELNDVLSQLPSRSSKITHRIHESASLRSBEET 1324
 Db 1258 KSTQDTWMALEALSEYWRIFKQDDNTLEIVNSLEKQSHQSKFRLKENLQEEIRSMGT 1317
 QY 1325 KENEGFTVTAEGKGGTSLVWVYHAKAKDQLTCNKFDILKVTIKPAPET----- 1373
 Db 1318 K----FNKIVSGKGGKILTVIKMYNLMQMTSCTELGLDVTVDACEAGWNEDDYDY 1373
 QY 1374 EKRQDQAKNWTMLICTRYRGDDAT-----MSTLIDISM 1407
 Db 1374 DEEMADPREPIVHDLRRSRREAFTEKEVKLLIYEVCLMKKSNVTLSGWALVDITL 1433
 QY 1408 MTGAPDDBLQIANGVDRYISKYELDKAFSDENTLIYLDKVSHEDDCLAFKHQYF 1467
 Db 1434 LSGIEPIDDLKLKLAESSEYISHYEYQ-----GRLLLYDFKVPDWT-DCVAFEAQTV 1487
 QY 1468 NVELIQGAVKVAAYVNLSESTFRFVHPEKEDKMLKCRDELRCRCAEBCFIQSDDKV 1527
 Db 1488 KVSLLQASAVIYDFEYFNKCTVFGAPSKPNFVTLCSGVQVCAEGLCPKLSPPKS 1547
 QY 1528 TLEBRDLKAC-BPGVDVYVYKTRLVKVLNSNDPEYIMAEQITKSGDS-VVVGQORTPI 1585
 Db 1548 TETERTQTFACYSPPRVFQVGVYVMENTAEQAQAFVYVTAIRLEVLQRNSDDIKPEPTRLFY 1607
 QY 1586 SPIKREALLKLEKHYLMGLSSDPWGEKPNLSYIIGKDTWVEHWEPEDECOEENOKQ 1645
 Db 1608 QPMCK-MRLAKETEYLINGQDQGVTKNEGHYILERKFWVEELPGEQKCAATRYNF 1665
 QY 1646 CODLGAFTESMVVFGC 1661
 Db 1666 CTDARNFMDKYKENGCG 1681

RESULT 14

CSHU
 Complement C5 precursor [validated] - human
 N:Contains: C5a anaphylatoxin; C5b
 C:Species: Homo sapiens (man)
 C:Date: 30-Sep-1992 #sequence revision 30-Sep-1992 #text_change 09-Jul-2004
 C:Accession: A40075; A27689; A01267; A01266; S15121
 R:Haviland, D.L.; Haviland, J.C.; Fleischer, D.T.; Hunt, A.; Wetsel, R.A.
 J. Immunol. 145, 362-368, 1991
 A:Title: Complete cDNA sequence of human complement pro-C5. Evidence of truncated transo
 A:Reference number: A40075; MUID:91079575; PMID:1984448
 A:Accession: A40075
 A:Molecule type: mRNA
 A:Residues: 1-1676 <HAV>
 A:Cross-references: UNIPROT:P01031; GB:M57729; NID:g179982; PIDN:AAA51925.1; PID:g179983
 A:Note: 518-Ser was also found
 R:Wetsel, R.A.; Lemons, R.S.; Le Beau, M.M.; Barnum, S.R.; Noack, D.; Tack, B.F.
 Biochemistry 27, 1474-1482, 1988
 A:Title: Molecular analysis of human complement component C5: localization of the struct
 A:Reference number: A27689; MUI-D:88209511; PMID:3365401
 A:Accession: A27689
 A:Molecule type: mRNA

A:Residues: 412-1676 <WET>
 A:Cross-references: GB:M65134; GB:M18879; NID:g179691; PIDN:AAA51856.1; PID:g179692
 R:Fernandez, H.N.; Hugli, T.E.
 J. Biol. Chem. 253, 6955-6964, 1978
 A:Title: Primary structural analysis of the polypeptide portion of human C5a anaphylato
 A:Reference number: A01267; MUID:79005687; PMID:690134
 A:Accession: A01267
 A:Molecule type: protein
 A:Residues: 678-751 <FER>
 R:Lundwall, A.B.; Wetsel, R.A.; Kristensen, T.; Whitehead, A.S.; Woods, D.E.; Ogden, R.
 J. Biol. Chem. 260, 2108-2112, 1985
 A:Title: Isolation and sequence analysis of a cDNA clone encoding the fifth complement
 A:Reference number: A01266; MUID:85103937; PMID:2579066
 A:Accession: A01266
 A:Molecule type: mRNA
 A:Residues: 412-854; S1ALSPRLCEKNGISCHCKLRIPGSSDSPASQVAGITGTHHAQPT' <LUN>
 A:Cross-references: GB:K02874
 A:Note: the carboxyl-terminal part of the sequence in this report appears to be derive
 R:Bohnack, J.F.; Mollison, K.W.; Buko, A.M.; Ashworth, J.C.; Hill, H.R.
 Biochem. J. 273, 635-640, 1991
 A:Title: Group B streptococci inactivate complement component C5a by enzymic cleavage a
 A:Reference number: S15121; MUID:91144547; PMID:1996961
 A:Contents: annotation
 C:Comment: Complement C5 contains two disulfide-linked chains, formed by removal of fou
 (beta and alpha' chains).
 C:Comment: Activation of C5 initiates the spontaneous assembly of the late complement c
 is the foundation upon which the membrane attack complex is assembled.
 C:Comment: C5a has potent spasmogenic and chemotactic activity.
 C:Genetics:
 A:Gene: GDB:C5
 A:Cross-references: GDB:119734; OMIM:120900
 A:Map position: 9q34.1-9q34.1
 C:Superfamily: alpha-2-macroglobulin
 C:Keywords: complement alternate pathway; complement pathway; cytolysis; glycoprotein;
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-673/678-1676/Product: complement C5 #status predicted <CSA>
 F:19-673/752-1676/Product: C5b #status predicted <CSB>
 F:19-673/Product: complement C5 and C5b beta chain #status predicted <CSBB>
 F:678-1676/Product: complement C5 alpha chain #status predicted <CSA>
 F:678-751/Product: C5a anaphylatoxin #status experimental <C5T>
 F:752-1676/Product: C5b alpha' chain #status predicted <CSBA>
 F:567-810/634-669,698-724,699-731,711-732,866-1527,1101-1159,1375-1505,1405-1474,1520-1
 F:741/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:751-752/Cleavage site: Arg-Leu (C5 convertase) #status experimental
 F:911,1115,1630/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 21.4%; Score 1840; DB 1; Length 1676;

Best Local Similarity 28.2%; Pred. No. 4.8e-104;

Matches 481; Conservative 343; Mismatches 759; Indels 124; Gaps 47;

QY 26 YSIITNLRLESEETMVALEHDAQDQVPTVTVTHDFPGKVLVLSSEKTVLPATNMGN 85
 Db 23 YVISAPKIFRVGASENIVIQVGYTEAFDATTISIKSYDPKFSYSSGHVLSSENKFNQS 82
 QY 86 VTFTIPANREFKSEKGRNKFVTQVTFQVVKVYVLSQSGVLFITDXTIYTPGSTV 145
 Db 83 AILTIQD-QQLPGGQNPVSVYVLEVSVKHSFKSKAMPYDNGFLFIHTDPVVTPOQSV 141
 QY 146 LYRIFTVNHKLIPVGTVMVNIENPEGIPVKQDSLSQNGVLPL-SWDIPELVNMGOW 204
 Db 142 KVAVSLNDLKPARETVLTIDEGSEV--DMVEEDHIGITISFPDFKIPSPRYGMW 199
 QY 205 KIRAYENSQQQVSTFEFVKVYVLPSPFEVIVPEKPYIYNEKGLVTTAFLYKCK 264
 Db 200 TIRAKYKEDFTTGTATYFVKVYVLPSPFEVIVPEKPYIYNEKGLVTTAFLYKCK 259
 QY 265 V-EGTAVFLFGQD---GQRIISLPESLKRIPIEDGSGEVLSRKVLKLDGQVNPRAEDLV 320
 Db 260 VTADVVIITFGREDLKDDQKEMQAMQNTWMLNGIAQVTFDSETAKELSYSLLEDLN 319
 QY 321 GKSLYSVATVILHSGSDMVAERSGIPVITSPYQIHFTKTPKYPKPGMPFOLMVFVTPD 380
 Db 320 NKLYIAVTVIESTGGFSEABIPGIKYVLSVKNLVATPLFLKPGIPYPIKQVKDSL 379

QY 381 GSPAYRVPVAVOGE-----DTVQSLQ-GDGVAKLISINTHPSQKPLSITVTRTK 428
Db 380 DLQVGGVPVILNAQTDVNNQTSDLDPKSVTRVDGVSFVLNLPSTVLEPNVKTDA 439
QY 429 QELSEAOATRTWQALPYSTVNSNNYLSLVRLTLPGETLVNFFLRMDRAHEAKIR 498
Db 440 PDLPEENQAREGYRAIAYSLSOSYLYIDWTDNHHKALLVGBHLMI--IVTPKSPYIDKIT 497
QY 489 YTYLWIKNGRLKAGQVREPGQDLVPLPSITDFIPSPRLVAYTYTLIGASQREWVA 548
Db 498 HYNVILSKGKLIHFGTREKFSASYSQINIPVTQNMVPPSKLLVLYIVITGEQ-TAELVS 556
QY 549 DSVWVDKSCVGLVVKVSGSEDRQPVGQOMTLKIEGDHAGARVLVADKGVFLNKK 608
Db 557 DSVWLVIEKCGNQLQVHLSPADAYS-PGQTVSLNMTGMSWVAALAAVDVAVYGVQRG 615
QY 609 NKLSTQKIWDVVEKADIGTGPCKDYAGVFSAGLTFTSSGGQQTQARAELOCPQPAAR 668
Db 616 AKKPLSERVFCLEKSLGCGAGGLNANVPHLAGLTFTLNANADDSQENDEPCKE-ILR 674
QY 669 RRSVQLTEKMDKV-GKYPKE-LAKCCEDGNRPNRPFSCQRTTRTSLGEACKVFLD 726
Db 675 PRRTLO---KKBEIAAKYKHSVVRKCCYDGCNVNDE-TCEQAAARLSLGPCKAFTE 730
QY 727 CCNYITELRRQARASHLGLARSNLD-EDIIAEENIVSRSEFPSSWLNWVBDLKEPPKNG 785
Db 731 CCVVASQLR---ANISHKDWQLGRHMKTLPLPVSKPEIRSYFPSSWLMWEHLV---PRR- 783
QY 786 ISYKLMNIFKDSITTWELAVMSDKGICVADPEVTVMQDFIDLRLPYSVVRNEOV 845
Db 784 ---KQLOFALDPLDITWEI--OGIGISN-TGICVADTVKAKVQKVFLEKWNIPYSVVRGEQI 839
QY 846 EIRAVLYNRQNLQELVRYELLHNPAFCSLATTKRRHQQTIT-----IPPKSSLSVPY 898
Db 840 QLKGTVNYR-TSGMGFCVKMSAVEGICTSESFVIDHQTKSKCVKRGVKGSSSHLVTF 898
QY 899 VIPLKTCQEVVEKAAVHHISDGVKSLVWPEGIRMKNTVAVRTLDERLGRGVQ 958
Db 899 TVLPLEIGHNINFSLEW--FGKEILVKTLLVWPEGVRSYSGV-TLDRPRGYIGISR 955
QY 959 KEDIP---PADLSDDQVPTSETRILLOQTPVAQMTEDAVDAERLKHILVTPSCGQNM 1015
Db 956 KXEPFVRIPDL---VPKTEIKRILSVKGLLVGELSAVLQSGEGINILTLHPKGSABEL 1012
QY 1016 IGMTPTVIAVHLYDETEQNEKF---GLEKRCQALELKKGYTQQLAFEPQPSAFAFVKR 1072
Db 1013 MSVVPVYVYHLETONHNI FHSDFLEKQKLKKLKEGMLSIMSYRNADYSYVWKG 1072
QY 1073 APSTWLTAYVVKVFLAVNLIAIDSQVLCGAVKMLLEKQKPDGVFQEDAPVHQMIGG 1132
Db 1073 SASTWLTAFALRVLGQVKNVYEQNONSICNSLLWLVENVQLDNGSFKENSQYQIKLOGT 1132
QY 1133 LR-NNNEKQALTAFLVLSIQAKDICEQVNSLPGSITKAGDFLEANYMLOBSYVAI 1191
Db 1133 LPVEARENSYLTAFVIGIRAFDIC--PLVKIDTALIKADNELLEWTLPAQSTFTLAI 1190
QY 1192 AGYALQMGRLKGLPLANKFLTAKOK-----NRWED-----PGKQLYN-----VE 1231
Db 1191 SAVALS-LGDKTHPQFESVSALKREALVKGNPPIYRFWKDNLQHKDSVNTGTARWE 1249
QY 1232 ATSVALLALQLKDFDPVVRVWRNEQRYGGVSGTQATFMVFOALAQOKDAPDHOE 1291
Db 1250 TTYALLTSLNLDINYNVNPVKNLSEQRVGGGYSTQDTINAIEGLTEVSLV---KK 1306
QY 1292 LNLDSVLQPSRSKTIHRIHWEASLL-RSEETKENEGFTV-TAEGKGQOTLSVWYTH 1349
Db 1307 LRLSMIDSVSKHKGALHYKMTKDNFLGRPEVLELNDLIVSTGFGSLGTLVHTVTVH 1366
QY 1350 AKAKDQLTGNKPLKTYIKPAETEKRPQDANKMTWILEICTRYRQDQATWS-----ILD 1404
Db 1367 KTSISEVCS-FYIKTDITQIEASHVRYGNSDYKRIIVACASYKPSRESSSGSHAVMD 1425

RESULT 15

CWS

Complement C5 precursor - mouse

N:Contains: C5a anaphylatoxin; C5b

C:Species: Mus musculus (house mouse)

C:Date: 19-Nov-1988 #sequence revision 15-Oct-1994 #text_change 09-Jul-2004

C:Accession: A35530; A27538; A40429

R:Wetzel, R.A.; Fleischer, D.T.; Haviland, D.L.

J. Biol. Chem. 265, 2435-2440, 1990

A:Title: Deficiency of the murine fifth complement component (C5). A 2-base pair gene de

A:Reference number: A35530; MUID:90153853; PMID:2303408

A:Accession: A35530

A:Molecule type: mRNA

A:Residues: 1-215; 'L' <WET>

A:Cross-references: UNIPROT:P06684; GB:M35526; GB:J05234; NID:g192302; PIDN:AAA37348.1;

R:Wetzel, R.A.; Ogata, R.T.; Tack, B.F.

Biochemistry 26, 737-743, 1987

A:Title: Primary structure of the fifth component of murine complement.

A:Reference number: A27538; MUID:87185363; PMID:2436653

A:Accession: A27538

A:Molecule type: mRNA

A:Residues: 'PGI', 44-1680 <WET2>

R:Haviland, D.L.; Haviland, J.C.; Fleischer, D.T.; Wetzel, R.A.

J. Biol. Chem. 266, 11818-11825, 1991

A:Title: Structure of the murine fifth complement component (C5) gene. A large, highly i

n: component genes.

A:Reference number: A40429; MUID:91268053; PMID:1711041

A:Accession: A40429

A:Molecule type: DNA

A:Residues: 1-15 <HAV>

A:Cross-references: GB:M64852

C:Comment: Complement C5 contains two disulfide-linked chains, formed by removal of four

(beta and alpha' chains).

C:Comment: Activation of C5 initiates the spontaneous assembly of the late complement cc

is the foundation upon which the membrane attack complex is assembled.

C:Comment: C5a has potent spasmogenic and chemotactic activity.

C:Genetics:

A:Map position: 2

A:Introns: 22/3; 86/3; 140/3; 164/3; 195/2; 223/1; 253/2; 291/3; 334/1; 372/3; 434/3; 51

3; 1224/1; 1292/3; 1343/3; 1364/3; 1392/1; 1411/2; 1445/3; 1470/3; 1506/1; 1534/1; 1564,

C:Superfamily: alpha-2-macroglobulin

F:19-674, 679-1679/Product: complement C5 #status predicted <C5b>

F:19-674, 756-1679/Product: C5b #status predicted <C5b>

F:19-674, 756-1679/Product: complement C5 and C5b beta chain #status predicted <C5b>

F:19-674, 756-1679/Product: complement C5 alpha chain #status predicted <C5a>

F:19-674, 756-1679/Product: C5a anaphylatoxin #status predicted <C5a>

F:19-674, 756-1679/Product: C5b alpha' chain #status predicted <C5b>

F:19-674, 756-1679/Product: C5b alpha' chain #status predicted <C5b>

F:19-674, 756-1679/Product: C5b alpha' chain #status predicted <C5b>

F:19-674, 756-1679/Product: C5b alpha' chain #status predicted <C5b>

F:19-674, 756-1679/Product: C5b alpha' chain #status predicted <C5b>

F:19-674, 756-1679/Product: C5b alpha' chain #status predicted <C5b>

F:19-674, 756-1679/Product: C5b alpha' chain #status predicted <C5b>

F:19-674, 756-1679/Product: C5b alpha' chain #status predicted <C5b>

F:19-674, 756-1679/Product: C5b alpha' chain #status predicted <C5b>

F:19-674, 756-1679/Product: C5b alpha' chain #status predicted <C5b>

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2004, 00:23:03 ; Search time 244 Seconds
(without alignments)
3921.508 Million cell updates/sec

Title: US-09-875-519A-22

Perfect score: 8609

Sequence: 1 MGPTSGPSLLLLTHPLA.....KQCQDLGARTSMVVGCPN 1663

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	8608	100.0	1663	2 AAR89906	Aar89906 homo sapi
2	8598	99.9	1663	1 CO3 HUMAN	P01024 homo sapien
3	6825	79.3	1663	1 CO3 RAT	P10126 rattus norv
4	6792.5	78.9	1666	1 CO3 CAVO	P12387 cavia porce
5	6750	78.4	1663	2 Q80X1	Q80X1 mus musculu
6	6749	78.4	1663	1 CO3 MOUSE	P01027 mus musculu
7	6724	78.1	1661	2 Q9GPF1	Q9GPF1 sus scrofa
8	4526.5	52.5	1652	2 Q90633	Q90633 gallus gall
9	4389.5	51.0	1651	1 CO3 NAJNA	Q01833 najja najja
10	4224	49.1	1642	2 Q91132	Q91132 najja kaouth
11	4206.5	48.9	1589	2 Q91588	Q91588 xenopus lae
12	3549	41.2	1642	2 Q9YTA9	Q9YTA9 cyprinus ca
13	3529	41.0	1642	2 Q9YIB0	Q9YIB0 cyprinus ca
14	3521	40.9	1645	2 Q9YTA7	Q9YTA7 cyprinus ca
15	3489	40.5	1640	2 Q9YTA8	Q9YTA8 cyprinus ca
16	3484.5	40.5	1640	1 CO3 ONCNY	P98093 oncorhynch
17	3368	39.1	1658	2 Q91BH1	Q91BH1 cryzias lat
18	3350.5	38.9	1677	2 Q6UQA6	Q6UQA6 ctenopharyn
19	3350.5	38.9	1677	2 AAQ74974	AAQ74974 ctenophar
20	3333	38.7	1655	2 Q9PY1	Q9PY1 paralichthy
21	3308.5	38.4	1657	2 Q9IBH0	Q9IBH0 cryzias lat
22	3254.5	37.8	1614	2 Q98977	Q98977 oncorhynch
23	3228.5	37.5	1684	2 Q9DDV9	Q9DDV9 oncorhynch
24	3205	37.2	1662	2 Q98TS6	Q98TS6 anarhchas
25	3082	35.8	726	1 CO3 RABIT	P12247 arctolagus
26	2632.5	30.6	1159	2 Q9YTA6	Q9YTA6 cyprinus ca
27	2526	29.3	904	2 Q9W61	Q9W61 prototenus
28	2490.5	28.9	1673	1 CO3 LAMUA	Q00685 lampetra ia
29	2438.5	28.3	1230	2 Q9YTA5	Q9YTA5 cyprinus ca
30	2148	25.0	1620	1 CO3 EPTBU	P98094 eptatretus
31	2122	24.6	1683	2 Q91741	Q91741 xenopus lae

32 2112 24.5 1693 2 Q76K14
33 2112 24.5 1693 2 BAC82347
34 2096.5 24.4 1700 2 Q9Y933
35 1913 22.2 1732 2 Q969A4
36 1857 21.6 1677 2 Q6YV1
37 1857 21.6 1677 2 AAQ97591
38 1848 21.5 1716 2 Q91932
39 1844.5 21.4 1724 2 Q70TF5
40 1844.5 21.4 1724 2 CAD66666
41 1840 21.4 1676 1 CO5 HUMAN
42 1817 21.1 1680 1 CO4 MOUSE
43 1815 21.1 1738 1 CO4 MOUSE
44 1815 21.1 1744 2 Q6UZE9
45 1815 21.1 1744 2 AAR89163

ALIGNMENTS

RESULT 1

AAR89906 PRELIMINARY; PRT; 1663 AA.
AC AAR89906;
DT 02-MAR-2004 (Tremblrel. 27, Created)
DT 02-MAR-2004 (Tremblrel. 27, Last sequence update)
DE Complement component 3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Carnivora; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Daniels R.L., da Ponte S.H., Hastings N.C., Ahearn M.O.,
RA Rajkumar N., Yi Q., Nickerson D.A.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY513239; AAR89906.1; -;
SQ SEQUENCE 1663 AA; 187147 MW; 30C2B32A9E75FFC4 CRC64;

Query Match 100.0%; Score 8608; DB 2; Length 1663;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1662; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGPTSGPSLLLLTHPLALGSPVSIITPILRLESEETVLEAHDAAQGDVPTVTVH 60
Db 1 MGPTSGPSLLLLTHPLALGSPVSIITPILRLESEETVLEAHDAAQGDVPTVTVH 50
Qy 61 DFGKLVLSSEKTVLTPATNMGNVTFTIPANREFKSEKGNKFTVQATFGTQWBEKV 120
Db 61 DFGKLVLSSEKTVLTPATNMGNVTFTIPANREFKSEKGNKFTVQATFGTQWBEKV 120
Qy 121 VLVSQSGYLFIQTDKTIYTPGSTVLYRFTVNHKLLPVGRVTVNVIENPEGIPVKQDSL 180
Db 121 VLVSQSGYLFIQTDKTIYTPGSTVLYRFTVNHKLLPVGRVTVNVIENPEGIPVKQDSL 180
Qy 181 SSQNQLGVPLSDWIDPELVNMQWKIRAYENSPOQVSTFEVKEYVLPSEFVIEPTE 240
Db 181 SSQNQLGVPLSDWIDPELVNMQWKIRAYENSPOQVSTFEVKEYVLPSEFVIEPTE 240
Qy 241 KFYIYNEXGLEVTITARELYGKVEGTAFVIFGIDGQRISLPESLKRIPIEDGSGEV 300
Db 241 KFYIYNEXGLEVTITARELYGKVEGTAFVIFGIDGQRISLPESLKRIPIEDGSGEV 300
Qy 301 VLRSKVLQGVQNPRAEDLVGKSLVYSATVILHSGSDMVQARSGIPITVSPYQIHFTKT 360
Db 301 VLRSKVLQGVQNPRAEDLVGKSLVYSATVILHSGSDMVQARSGIPITVSPYQIHFTKT 360
Qy 361 PKYFKGMPFDLMVFTNPDGSPAYRVPVAVOGEDTVQSLTQGDGVAKLINTHPSQKPL 420
Db 361 PKYFKGMPFDLMVFTNPDGSPAYRVPVAVOGEDTVQSLTQGDGVAKLINTHPSQKPL 420
Qy 421 SITVTRKKLSLSAEQATRTMQALPYSTVGNSSNYLHLVLRTELPRQGETLVNAPFLAMD 480

Db 421 SITVTRKKQELSEACATRTYQALPYSTVGNSSNNYHLISVLRLTELPGETLVNVLFLRMD 480
Qy 481 RAHEAKIRYYTYLIMNKGHLLKAGQVREPGQDLVVLPLSLTTDFIPFSRLVAYTYLIGA 540
Db 481 RAHEAKIRYYTYLIMNKGHLLKAGQVREPGQDLVVLPLSLTTDFIPFSRLVAYTYLIGA 540
Qy 541 SGQREVAVSVVDVKDSVGLSVKSGQSDRQVPGQOMTLKIEGDHGVAVLVAVDK 600
Db 541 SGQREVAVSVVDVKDSVGLSVKSGQSDRQVPGQOMTLKIEGDHGVAVLVAVDK 600
Qy 601 GVFVNLKKNLTKOSKIWDVWEKADIGCTPGSGKDVAGVPSDAGLTFTSSSCQQAQRAEL 660
Db 601 GVFVNLKKNLTKOSKIWDVWEKADIGCTPGSGKDVAGVPSDAGLTFTSSSCQQAQRAEL 660
Qy 661 QCQPQARRRRSQQTEKMDVGVKPKELKCCEDGKREMPNPFSCORRTFFSLGECAC 720
Db 661 QCQPQARRRRSQQTEKMDVGVKPKELKCCEDGKREMPNPFSCORRTFFSLGECAC 720
Qy 721 KKVFLDCCNVITELRQHARASHGLGARSNLDEDIIAENIVSRSEPPESHLWVVDLKE 780
Db 721 KKVFLDCCNVITELRQHARASHGLGARSNLDEDIIAENIVSRSEPPESHLWVVDLKE 780
Qy 781 PKXNGISTKLNNIFLKDSTITWEILLAVSMDSKGIQVADPFEVTVMQDFFIDLRPLYSV 840
Db 781 PKXNGISTKLNNIFLKDSTITWEILLAVSMDSKGIQVADPFEVTVMQDFFIDLRPLYSV 840
Qy 841 RNEQVEIRAVLYNQRELKVRVELLNHPAFCSLATTYKRHOQTITIPKSSLSVPIVI 900
Db 841 RNEQVEIRAVLYNQRELKVRVELLNHPAFCSLATTYKRHOQTITIPKSSLSVPIVI 900
Qy 901 VPLKGTGLQEVKAAVYHFIIDGVRKSLKVYVPEGRMKNKTAVRVLDPERLGRGVOKE 960
Db 901 VPLKGTGLQEVKAAVYHFIIDGVRKSLKVYVPEGRMKNKTAVRVLDPERLGRGVOKE 960
Qy 961 DIPADLSQVPTSETSEIILQGTVPVQAMTDAVDAERLKLIVTPSGCGEQNMIGWP 1020
Db 961 DIPADLSQVPTSETSEIILQGTVPVQAMTDAVDAERLKLIVTPSGCGEQNMIGWP 1020
Qy 1021 TVIAHYLDTEQWERFGLKQKQALELTKGTYTQQLAFROPSSAFAPVFRAPSTWLTA 1080
Db 1021 TVIAHYLDTEQWERFGLKQKQALELTKGTYTQQLAFROPSSAFAPVFRAPSTWLTA 1080
Qy 1081 YVVKVPSLANVLAISQVLCGAVKWLILKQKPGVFOEDAPVYHQEMICGLRNNKED 1140
Db 1081 YVVKVPSLANVLAISQVLCGAVKWLILKQKPGVFOEDAPVYHQEMICGLRNNKED 1140
Qy 1141 MALTAFLVLSIQBAKDCIEQVNSLPGSITKAGDFLEANYMNLQRSYTVATAGYALQMG 1200
Db 1141 MALTAFLVLSIQBAKDCIEQVNSLPGSITKAGDFLEANYMNLQRSYTVATAGYALQMG 1200
Qy 1201 RLKGPLLNKELTTAKQKNWEDPGKOLYVNEATSVALLALQLKDFDVPVFWLNEOR 1260
Db 1201 RLKGPLLNKELTTAKQKNWEDPGKOLYVNEATSVALLALQLKDFDVPVFWLNEOR 1260
Qy 1261 YVGGYGSTQATFWVQALAQYQKADPQHCELNLDVSLQPSRSKITHRIHWESASILLR 1320
Db 1261 YVGGYGSTQATFWVQALAQYQKADPQHCELNLDVSLQPSRSKITHRIHWESASILLR 1320
Qy 1321 SEETKENEGTVAEGKGGTISVYTMVYHAKAKDQLTCKNEDLKVITKPAETEKRPQDA 1380
Db 1321 SEETKENEGTVAEGKGGTISVYTMVYHAKAKDQLTCKNEDLKVITKPAETEKRPQDA 1380
Qy 1381 KNTWILEICRYRGDQDATMSILDISMTMGFAPDITDQLKQANGVDRYISKYLEDKAFSD 1440
Db 1381 KNTWILEICRYRGDQDATMSILDISMTMGFAPDITDQLKQANGVDRYISKYLEDKAFSD 1440
Qy 1441 RNTLIIIVLDKVSHEDDCLAFKHQYVNVLTQPGAVKVAYNLESCITFPYHPEKDG 1500
Db 1441 RNTLIIIVLDKVSHEDDCLAFKHQYVNVLTQPGAVKVAYNLESCITFPYHPEKDG 1500
Qy 1501 KLNKLCRDELRCRAEENCFIQKSDDKVTLEERLDKACBFGVDVYVYKTLRVQLSNDFDE 1560
Db 1501 KLNKLCRDELRCRAEENCFIQKSDDKVTLEERLDKACBFGVDVYVYKTLRVQLSNDFDE 1560

RESULT 2

CO3_HUMAN
ID CO3_HUMAN STANDARD; PRT; 1663 AA.
AC P01024;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Complement C3 precursor [Contains: C3a anaphylatoxin].
GN Name=C3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85140166; PubMed=2579379;
RA de Bruijn M.H.L., Fey G.H.;
RT "Human complement component C3: cDNA coding sequence and derived
primary structure.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:708-712 (1985).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANTS GLY-102; PRO-314; LYS-863; ASP-1224
AND THR-1367.
RA Bieder M.J., Daniels R.L., da Ponte S.H., Hastings N.C., Ahearn M.O.,
RA Rajkumar N., Yi Q., Nickerson D.A.;
RT "SeattlesNPS, NHUHI HLB682 program for genomic applications, UW-
PHCRC, Seattle, WA (URL: http://pga.gs.washington.edu).";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 672-748.
RX MEDLINE=76069169; PubMed=1238393;
RA Hugli T.E.;
RT "Human anaphylatoxin (C3a) from the third component of complement.
Primary structure.";
RL J. Biol. Chem. 250:8293-8301 (1975).
RN [4]
RP SEQUENCE OF 955-966, AND SUBUNITS.
RC TISSUE=Serum;
RX MEDLINE=95293954; PubMed=7539791;
RA Oxvig C., Haaning J., Kristensen L., Wagner J.M., Rubin I.,
RA Stigbrand T., Gleich G.J., Sottrup-Jensen L.;
RT "Identification of angiotensinogen and complement C3dg as novel
protein binding the proform of eosinophil major basic protein in
human pregnancy serum and plasma.";
RL J. Biol. Chem. 270:13645-13651 (1995).
RN [5]
RP SEQUENCE OF 988-1036.
RX MEDLINE=82174534; PubMed=6175959;
RA Thomas M.L., Janatova J., Gray W.R., Tack B.F.;
RT "Third component of human complement: localization of the internal
thiolester bond.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:1054-1058 (1982).
RN [6]
RP SEQUENCE OF 1409-1563.
RX MEDLINE=88154452; PubMed=3279119;
RA Daoudaki M.E., Becherer J.D., Lambiris J.D.;
RT "A 34-amino acid peptide of the third component of complement mediates
properdin binding.";
RL J. Immunol. 140:1577-1580 (1988).
RN [7]
RP STRUCTURE BY NMR OF C3A.
RX MEDLINE=88276894; PubMed=3260670;
RA Nettesheim D.G., Edalji R.P., Mollison K.W., Greer J.,

RA Zuideweg B.R.P.;
 RT "Secondary structure of complement component C3a anaphylatoxin in
 RT solution as determined by NMR spectroscopy: differences between
 RT crystal and solution conformations.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:5036-5040(1988).
 [8]
 RN MUTAGENESIS OF THIOESTER BOND REGION.
 RX MEDLINE=92250565; PubMed=157777;
 RA Isaac L., Isenman D.E.;
 RT "Structural requirements for thioester bond formation in human
 RT complement component C3. Reassessment of the role of thioester bond
 RT integrity on the conformation of C3.";
 RL J. Biol. Chem. 267:10062-10069(1992).
 [9]
 RN DISULFIDE BONDS.
 RX MEDLINE=93106233; PubMed=8416818;
 RA Dolmer K., Sottrup-Jensen L.;
 RT "Disulfide bridges in human complement component C3b.";
 RL FEBS Lett. 315:85-90(1993).
 [10]
 RN CARBOHYDRATE-LINKAGE SITE ASN-85.
 RX MEDLINE=22660472; PubMed=12754519; DOI=10.1038/nbt827;
 RA Zhang H., Li X.-J., Martin D.B., Aebersold R.;
 RT "Identification and quantification of N-linked glycoproteins using
 RT hydrazide chemistry: stable isotope labeling and mass spectrometry.";
 RL Nat. Biotechnol. 21:660-666(2003).
 [11]
 RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 996-1303.
 RX MEDLINE=98259089; PubMed=9596584;
 RA Nagar B., Jones R.G., Diefenbach R.J., Isenman D.E., Rini J.M.;
 RT "X-ray crystal structure of C3d: a C3 fragment and ligand for
 RT complement receptor 2.";
 RL Science 280:1277-1281(1998).
 [12]
 RN VARIANT C3F/S.
 RP MEDLINE=89309808; PubMed=2473125;
 RA Poznansky M.C., Clissold P.M., Lachmann P.J.;
 RT "The difference between human C3F and C3S results from a single amino
 RT acid change from an asparagine to an aspartate residue at position
 RT 1216 on the alpha-chain of the complement component, C3.";
 RL J. Immunol. 143:1254-1258(1989).
 [13]
 RN ERRATUM (RETRACTION).
 RX MEDLINE=90063087; PubMed=2584723;
 RA Poznansky M.C., Clissold P.M., Lachmann P.J.;
 RL J. Immunol. 143:3860-3862(1989).
 [14]
 RN VARIANTS GLY-102 AND PRO-314.
 RX MEDLINE=91011240; PubMed=1976733;
 RA Boto M., Yong Fong K., So A.K., Koch C., Walport M.J.;
 RT "Molecular basis of polymorphisms of human complement component C3.";
 RL J. Exp. Med. 172:1011-1017(1990).
 [15]
 RN VARIANT ASN-549.
 RX MEDLINE=95050640; PubMed=7961791;
 RA Singer L., Whitehead W.T., Akama H., Katz Y., Fishelson Z.,
 RA Wiesel R.A.;
 RT "Inherited human complement C3 deficiency. An amino acid substitution
 RT in the beta-chain (Asp549 to Asn) impairs C3 secretion.";
 RL J. Biol. Chem. 269:28494-28499(1994).
 [16]
 RN VARIANT GLN-1320.
 RA Watanabe Y., Matsui N., Yan K., Nishimukai H., Tokunaga K., Juji T.,
 RA Kobayashi N., Kohsaka T.;
 RT "A novel C3 allotype C3'F02' has an amino acid substitution that may
 RT inhibit iC3b synthesis and cause C3-hypocomplementemia.";
 RL Mol. Immunol. 30:62-62(1993).
 CC -!- FUNCTION: C3 plays a central role in the activation of the
 CC complement system. Its processing by C3 convertase is the central
 CC reaction in both classical and alternative complement pathways.
 CC After activation C3b can bind covalently, via its reactive
 CC thioester, to cell surface carbohydrates or immune aggregates.
 CC -!- FUNCTION: Derived from proteolytic degradation of complement C3,

CC C3a anaphylatoxin is a mediator of local inflammatory process. It
 CC induces the contraction of smooth muscle, increases vascular
 CC permeability and causes histamine release from mast cells and
 CC basophilic leukocytes.
 CC -!- SUBUNIT: C3 precursor is first processed by the removal of 4 Arg
 CC residues, forming two chains, beta and alpha, linked by a
 CC disulfide bond. C3 convertase activates C3 by cleaving the alpha
 CC chain, releasing C3a anaphylatoxin and generating C3b (beta chain
 CC + alpha chain). During pregnancy, C3dg exists as a complex
 CC (probably a 2:2:2 heterohexamer) with AGR and the proform of PRG2.
 CC -!- PM: C3b is rapidly split in two positions by factor I and a
 CC cofactor to form iC3b (inactivated C3b) and C3f which is released.
 CC Then iC3b is slowly cleaved (possibly by factor I) to form C3c and
 CC C3dg. Other proteases produce other fragments such as C3d or C3g.
 CC -!- POLYMORPHISM: There are two alleles: C3S (C3 slow), the most
 CC common allele in all races and C3F (C3 fast), relatively frequent
 CC in Caucasoids, less common in Black Americans, extremely rare in
 CC Orientals.
 CC -!- DISEASE: Defects in C3 are the cause of C3 deficiency
 CC (MIM:120700). It can result in susceptibility to pyogenic
 CC infection.
 CC -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
 CC -!- SIMILARITY: Contains 1 NTR domain.
 CC ---
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; K02765; AAA85332.1; -;
 DR EMBL; AY513239; AAR89906.1; -;
 DR PIR; A94065; C3HU.
 DR PDB; 1C3D; X-ray; @=-;
 DR PDB; 1GHQ; X-ray; A=994-1300.
 DR SWISS-2DPAGE; P01024; HUMAN.
 DR Sienna-2DPAGE; P01024; -;
 DR Genew; HGNC:1318; C3.
 DR MIM; 120700; -;
 DR GO; GO:0005102; F:receptor binding; TAS.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.
 DR GO; GO:0006955; P:immune response; TAS.
 DR GO; GO:0007165; P:signal transduction; TAS.
 DR InterPro; IPR002890; A2M_N.
 DR InterPro; IPR009048; AM_Receptor_bind.
 DR InterPro; IPR000020; Anaphylatoxin.
 DR InterPro; IPR01840; Anaphylatoxin.
 DR InterPro; IPR008964; Invasin_intimin.
 DR InterPro; IPR01599; MacroglobulinA2.
 DR InterPro; IPR01134; Netrin_C.
 DR InterPro; IPR008930; Terp_cyc_toroid.
 DR InterPro; IPR008993; TIMP_like.
 DR Pfam; PF00207; A2M; 1.
 DR Pfam; PF01835; A2M_N; 1.
 DR Pfam; PF01821; ANATO; 1.
 DR Pfam; PF01759; NTR; 1.
 DR PRINTS; PR00004; ANAPHYLATOXN.
 DR ProDom; PD003264; Anaphylatoxin; 1.
 DR PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
 DR PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
 DR PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
 DR PROSITE; PS00189; NTR; 1.
 KW 3D-structure; Complement alternate pathway; Complement pathway;
 KW Direct protein sequencing; Disease mutation; Glycoprotein;
 Query Match 99.9%; Score 8598; DB 1; Length 1663;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1661; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MGSTGSPSLLLLLTHPLALGSPMSYITNPLSEFTWLEHDAQGVVTVVH 60
 |||||

Db 1 MGTSGPSLLLLTHLPALGSPMYSIITPNILBLESEETWVLEAHDAQDVPVTVVH 60
Qy 61 DFGKKLVLSSEKTVLTPATNMGVNTFTIPANREFKSEKGRNFVTVQATFGQVVEKV 120
Db 61 DFGKKLVLSSEKTVLTPATNMGVNTFTIPANREFKSEKGRNFVTVQATFGQVVEKV 120
Qy 121 VLVSQSGYLFQTDKXTIYTPGSTVLYRIFTVNHKLLPVGTVMVNIENPEGIPVKQDSL 180
Db 121 VLVSQSGYLFQTDKXTIYTPGSTVLYRIFTVNHKLLPVGTVMVNIENPEGIPVKQDSL 180
Qy 181 SSQNLGVLPLSDIPELVNMGQWKIRAYENSPQVFEFVEKVEYVLPSEFIVEPTE 240
Db 181 SSQNLGVLPLSDIPELVNMGQWKIRAYENSPQVFEFVEKVEYVLPSEFIVEPTE 240
Qy 241 KFYIYNEKGLVITITARELYKKGVEGTAFLVFGQDGEQRIISLPESIKRIPIEDGSGEV 300
Db 241 KFYIYNEKGLVITITARELYKKGVEGTAFLVFGQDGEQRIISLPESIKRIPIEDGSGEV 300
Qy 301 VLSRKVLGQVQNPRAEDLVGSLYVSATVILHSGSDMVQAERSGIPITVSPYQIHFTKT 360
Db 301 VLSRKVLGQVQNPRAEDLVGSLYVSATVILHSGSDMVQAERSGIPITVSPYQIHFTKT 360
Qy 361 PKYFKGMPDLMVFTNPDGSPAYRVPVAVOGEDTVQSLTQGDGVAKLSINTHPSOKPL 420
Db 361 PKYFKGMPDLMVFTNPDGSPAYRVPVAVOGEDTVQSLTQGDGVAKLSINTHPSOKPL 420
Qy 421 SITVRTKQELSEAEQATRTMOALPYSTVGNSSNNYLHLSVLRTLPGETLNVNFLRMD 480
Db 421 SITVRTKQELSEAEQATRTMOALPYSTVGNSSNNYLHLSVLRTLPGETLNVNFLRMD 480
Qy 481 RAHEAKIRYTYLIMNKGKLLKAGQVREPQDVLVPLSITTFIPFSFRVAVYTLIGA 540
Db 481 RAHEAKIRYTYLIMNKGKLLKAGQVREPQDVLVPLSITTFIPFSFRVAVYTLIGA 540
Qy 541 SGQREVYADSVVDVKDSVCSVLVKSQSDQRPVPCQMTLKIEGDHGARVVLVAVDK 600
Db 541 SGQREVYADSVVDVKDSVCSVLVKSQSDQRPVPCQMTLKIEGDHGARVVLVAVDK 600
Qy 601 GVFLVANKKLTOSKTDWVEKADIGCTPGSGKDVAGVPSDAGLTFSTSSSQQTAAQRAEL 660
Db 601 GVFLVANKKLTOSKTDWVEKADIGCTPGSGKDVAGVPSDAGLTFSTSSSQQTAAQRAEL 660
Qy 661 QCPQPAARRRSVQLTEKMDKVGYPKELRKCCEGDMRENPMRPFSCQRRTRFISLGEAC 720
Db 661 QCPQPAARRRSVQLTEKMDKVGYPKELRKCCEGDMRENPMRPFSCQRRTRFISLGEAC 720
Qy 721 KKVFLDCCNVITELRRQHARASHGLGARSNLDEIIAENIVSRSEFPESLWNVDELKE 780
Db 721 KKVFLDCCNVITELRRQHARASHGLGARSNLDEIIAENIVSRSEFPESLWNVDELKE 780
Qy 781 PPKNGISTKLWNITFLKDSITTWELIAVMSDKKGCIVADPFEVTVMQDFFIDLRPLSVV 840
Db 781 PPKNGISTKLWNITFLKDSITTWELIAVMSDKKGCIVADPFEVTVMQDFFIDLRPLSVV 840
Qy 841 RNQVEIRAVLYNRQNELKVRVELLNHPACSLATTKRHOQITIPPKSSLSVPYVI 900
Db 841 RNQVEIRAVLYNRQNELKVRVELLNHPACSLATTKRHOQITIPPKSSLSVPYVI 900
Qy 901 VPLKTLGQVEVZAAVYHHFISDGVKSLKVPVEGIRMNKTVAVRTLDPERLREGVQKE 960
Db 901 VPLKTLGQVEVZAAVYHHFISDGVKSLKVPVEGIRMNKTVAVRTLDPERLREGVQKE 960
Qy 961 DIPPADLSDQVPTSETRIILQGTTPVAQMTEDAVDAERLKHILVTPSGCGEQNWGMT 1020
Db 961 DIPPADLSDQVPTSETRIILQGTTPVAQMTEDAVDAERLKHILVTPSGCGEQNWGMT 1020
Qy 1021 TVTAVHVLDETEBQWFKGLKROGALELTKGYTQOLAPRQSSAFAPFVGRAPSTWLTA 1080
Db 1021 TVTAVHVLDETEBQWFKGLKROGALELTKGYTQOLAPRQSSAFAPFVGRAPSTWLTA 1080
Qy 1081 YVVKVFLAVNLTAIDSQVLCGAVKWLILEKQKPGVFOEDAPVHQEMIGGLNNNEKD 1140
Db 1081 YVVKVFLAVNLTAIDSQVLCGAVKWLILEKQKPGVFOEDAPVHQEMIGGLNNNEKD 1140

Qy 1141 MALTAFLVLSLOBAKDI CEEQVNSLPGSITKAGDFLEANNMQLORSYTVTAIAGYALAOQG 1200
Db 1141 MALTAFLVLSLOBAKDI CEEQVNSLPGSITKAGDFLEANNMQLORSYTVTAIAGYALAOQG 1200
Qy 1201 RLKGPLLNLKFLTTAKDKVRWEDPGKQLYVVEATSYALLALLQKDFDFPVPVVRWLNQOR 1260
Db 1201 RLKGPLLNLKFLTTAKDKVRWEDPGKQLYVVEATSYALLALLQKDFDFPVPVVRWLNQOR 1260
Qy 1261 YGGYGSTQATFVMPVFOALAOYKQADPHQELNLDVSLQPSRSSKITHRIHWESASLLR 1320
Db 1261 YGGYGSTQATFVMPVFOALAOYKQADPHQELNLDVSLQPSRSSKITHRIHWESASLLR 1320
Qy 1321 SEETKEGEGFTVTAEGKGQGTLSVVTMYHAKAKDQLTCKNFKDKVTIKPAPETERPQDA 1380
Db 1321 SEETKEGEGFTVTAEGKGQGTLSVVTMYHAKAKDQLTCKNFKDKVTIKPAPETERPQDA 1380
Qy 1381 KNTWILBICTRYRGDQDATWSIILDISMTGFAEDTDQLKQLANGVDRIYSKYELDKAFSD 1440
Db 1381 KNTWILBICTRYRGDQDATWSIILDISMTGFAEDTDQLKQLANGVDRIYSKYELDKAFSD 1440
Qy 1441 RNTLIIVLDKVSHSEDDCLAFKVHQYFNVELIQFGAVKYVAYYNLEESCTREFYHPEKEDG 1500
Db 1441 RNTLIIVLDKVSHSEDDCLAFKVHQYFNVELIQFGAVKYVAYYNLEESCTREFYHPEKEDG 1500
Qy 1501 KLNKCRDELRCRAEENCFIQKSDDKVTLEERLDKACERGVYVYKTRLVKQLSNDPDE 1560
Db 1501 KLNKCRDELRCRAEENCFIQKSDDKVTLEERLDKACERGVYVYKTRLVKQLSNDPDE 1560
Qy 1561 YMAIEQITIKSGSDEVQVQQRFTFISPIKCREALKLEEKHYLMGLSSDFWGEKPNLSY 1620
Db 1561 YMAIEQITIKSGSDEVQVQQRFTFISPIKCREALKLEEKHYLMGLSSDFWGEKPNLSY 1620
Qy 1621 IIGKDTWVHRWPEDEDEQDEENKQCODIGAFPTESMVVFGCPN 1663
Db 1621 IIGKDTWVHRWPEDEDEQDEENKQCODIGAFPTESMVVFGCPN 1663

RESULT 3
CO3_RAT STANDARD; PRT; 1663 AA.
ID CO3_RAT
AC P01026;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Complement C3 precursor [Contains: C3a anaphylatoxin].
GN Name=C3;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mistar; TISSUE=Liver;
RX MEDLINE=90245672; PubMed=2336397;
RA Misumi Y., Sohma M., Ikehara Y.;
RT "Nucleotide and deduced amino acid sequence of rat complement C3.";
RL Nucleic Acids Res. 18:2178-2178(1990).
RN [2]
RP SEQUENCE OF 671-748.
RX MEDLINE=79062262; PubMed=309768;
RA Jacobs J.W., Rubin J.S., Hugli T.E., Bogardt R.A., Mariz I.K.,
RA Daniels J.S., Daughaday W.H., Bradshaw R.A.;
RT "Purification, characterization, and amino acid sequence of rat anaphylatoxin (C3a).";
RL Biochemistry 17:5031-5038 (1978).
RN [3]
RP SEQUENCE OF 1316-1595 FROM N.A.
RX MEDLINE=89380332; PubMed=2674144;
RA Sundstrom S.A., Komm B.S., Ponce-De-Leon H., Yi Z., Teuscher C.,
RA Lyttle C.R.;
RT "Estrogen regulation of tissue-specific expression of complement C3.";
RL J. Biol. Chem. 264:16941-16947(1989).

CC -|- FUNCTION: C3 plays a central role in the activation of the complement system. Its processing by C3 convertase is the central reaction in both classical and alternative complement pathways. After activation C3b can bind covalently, via its reactive thioester, to cell surface carbohydrates or immune aggregates. C3a anaphylatoxin is a mediator of local inflammatory process. It induces the contraction of smooth muscle, increases vascular permeability and causes histamine release from mast cells and basophilic leukocytes.

CC -|- SUBUNIT: C3 precursor is first processed by the removal of 4 Arg residues, forming two chains, beta and alpha, linked by a disulfide bond. C3 convertase activates C3 by cleaving the alpha chain, releasing C3a anaphylatoxin and generating C3b (beta chain + alpha' chain).

CC -|- SIMILARITY: Contains 1 anaphylatoxin-like domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; X52477; CAA36716.1; -

DR EMBL; M29866; AAB40837.1; ALT_SEQ.

DR PIR; S15764; C3RT.

DR PDB; 1QQF; X-ray; A=1010-1286.

DR PDB; 1OSG; X-ray; A/B/C/D=1010-1286.

DR RGD; 2232; C3.

DR InterPro; IPR002890; A2M_N.

DR InterPro; IPR009048; AM_Receiver_bind.

DR InterPro; IPR000020; Anaphylatoxin.

DR InterPro; IPR001840; Anaphylatoxin.

DR InterPro; IPR008964; Invasin_intimin.

DR InterPro; IPR001599; MacroglobulinA2.

DR InterPro; IPR001134; Netrin_C.

DR InterPro; IPR008930; Terp_cyc_toroid.

DR InterPro; IPR008993; TIMP_like.

DR Pfam; PF00207; A2M; 1.

DR Pfam; PF01835; A2M_N; 1.

DR Pfam; PF01821; ANATO; 1.

DR Pfam; PF01753; NTR; 1.

DR PRINTS; PR000004; ANAPHYLATOXN.

DR PRODOM; PD03264; Anaphylatoxin; 1.

DR PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.

DR PROSITE; PS01177; ANAPHYLATOXIN_1; 1.

DR PROSITE; PS01178; ANAPHYLATOXIN_2; 1.

DR PROSITE; PS0189; NTR; 1.

KW 3D-structure. Complement alternate pathway; Complement pathway;

KW Direct protein sequencing; Glycoprotein; Inflammatory response;

KW Plasma; Signal; Thioester bond.

FT SIGNAL 1 24

FT CHAIN 25 1663 Complement C3.

FT CHAIN 25 666 Complement C3 beta chain.

FT CHAIN 671 1663 Complement C3 alpha chain.

FT PEPTIDE 671 748 C3a anaphylatoxin.

FT CHAIN 749 1663 Complement C3b alpha' chain.

FT DOMAIN 693 728 Anaphylatoxin-like.

FT DOMAIN 1518 1661 NTR.

FT SITE 748 749 Cleavage (by C3 convertase).

FT DISULFID 558 816 Interchain (by similarity).

FT DISULFID 626 661 By similarity.

FT DISULFID 693 720 By similarity.

FT DISULFID 694 727 By similarity.

FT DISULFID 707 728 By similarity.

FT DISULFID 873 1513 By similarity.

FT DISULFID 1101 1158 By similarity.

FT DISULFID 1358 1489 By similarity.

FT DISULFID 1389 1458 By similarity.

FT DISULFID 1506 1511 By similarity.

FT DISULFID 1518 1590 By similarity.

FT DISULFID 1537 1661 By similarity.

FT CROSSLINK 1010 1013 N-linked (GlcNAc...) (Probable).

FT CARBOHYD 939 939 N-linked (GlcNAc...) (Probable).

FT CARBOHYD 1617 1617 N-linked (GlcNAc...) (Probable).

FT CONFLICT 721 722 LX -> KL (in Ref. 2).

FT TURN 1011 1012

FT HELIX 1013 1031

FT TURN 1032 1032

FT TURN 1034 1037

FT HELIX 1039 1041

FT HELIX 1042 1057

FT TURN 1058 1059

FT STRAND 1060 1060

FT TURN 1062 1063

FT STRAND 1066 1066

FT TURN 1070 1071

FT HELIX 1076 1089

FT TURN 1090 1092

FT HELIX 1097 1111

FT STRAND 1112 1112

FT TURN 1114 1115

FT STRAND 1118 1118

FT HELIX 1127 1134

FT TURN 1137 1138

FT HELIX 1139 1158

FT TURN 1159 1161

FT TURN 1163 1164

FT HELIX 1165 1180

FT TURN 1181 1182

FT HELIX 1186 1198

FT TURN 1199 1200

FT TURN 1204 1205

FT HELIX 1206 1213

FT STRAND 1215 1215

FT TURN 1216 1218

FT STRAND 1219 1219

FT TURN 1223 1224

FT HELIX 1226 1242

FT TURN 1243 1244

FT TURN 1246 1247

FT HELIX 1249 1258

FT TURN 1263 1264

FT TURN 1266 1267

FT HELIX 1269 1285

SEQ SEQUENCE 1663 AA; 186460 MW; 2F87CCB143CDD4BC CRC64;

Query Match 79.3%; Score 6825; DB 1; Length 1663;

Best Local Similarity 78.1%; Pred. No. 0;

Matches 1301; Conservative 171; Mismatches 188; Indels 6; Gaps 5;

Qy 1 MGTSQPSL--LLLLTHLPALGSPMYSITTPNLRLESEETVLEAHDAGQDPVTVT 58

Db 1 MGFTSGQLLVLLLLSLALGSPMYSITTPNLRLESEETVLEAHDAGQDPVTVT 60

Qy 59 VHDFFGKLVLSSEKTVLTPATNMGVNTFTIPANREFKSEKRNKFTVTCATGTOVVE 118

Db 61 VQDFL-KKQVLTSEKTVLTGATLNRVFIKIPASKEFNADKG-HKYVTVVANGATVVE 118

Qy 119 KVVVLSQSGYLPIDTKTITTPGTVLYRFTVNNHKLTPVGRVNMNIENPEGIPKQD 178

Db 119 KAVLSVFSQSGYLPIDTKTITTPGTVLYRFTVNNHKLTPVGRVNMNIENPEGIPKRD 178

Qy 179 SLSSQNLQGLVPLSWDIPELVNMGOWKIRAYVENSPOQVFTSEFEVKEYLPSFEVIVEP 238

Db 179 ILSSHNQYGLPLSNWIPELVNMGOWKIRAFVHAPKQTFSAEFVKEYLPSFEVIVEP 238

Qy 239 TKFYIYINKEGLEVTITRFLYKKGVEGTAFVIFGIQGGQRISLPESLKRIEDGSG 298

Db 239 TKFYIHGPKGLEVSITARFLYKGNVDGTAFFVPGVQDEDKISLALSTRVIEDGSG 298

Qy 299 EYVLSKVLDDGVQNPRAEDLVGKSLYVSATVILHSGSDMVQASRGIPVTSYQIHFT 358

Db 299 EAVLSRKVLMDGVRPSSPALVGKSLVSVTVILHSGSDMVEABSGIPVITSPVQIHFT 358
Qy 359 KTKPKYKPGVDFDLVFTNPDGSPAYRVPVAVQGEDTVQSLTQDGDGAKLSINTHPSQK 418
Db 359 KTKPKYKPGVDFDLVFTNPDGSPAYRVPVAVQGEDTVQSLTQDGDGAKLSINTHPSQK 417
Qy 419 PLSTVTRKQELSEABQATRTMQALPYSTGNSNNYLHLSVLRTPELPGSTLNVFLLR 478
Db 418 PLSTVTRKQELSEABQATRTMQALPYSTGNSNNYLHLSVLRTPELPGSTLNVFLLR 477
Qy 479 MDRAHEAKIRYVTVLIMNKGRLKAGQVREPGQDLVPLISITTFPSRLVAVYTLI 538
Db 478 TDAGEAKIRYVTVLIMNKGRLKAGQVREPGQDLVPLISITTFPSRLVAVYTLI 537
Qy 539 GASQOREVADSVWVDKSGVSLVVKSGQSEDEQPVPGQMTLKISGDHARVLVAV 598
Db 538 GASQOREVADSVWVDKSGVSLVVKSGQSEDEQPVPGQMTLKISGDHARVLVAV 597
Qy 599 DKGVLVKNKXKLTQSKIWDVVEKADIGCTPGSGDKYAGVPSDAGLTFTSSGGQTAQRA 658
Db 598 DKGVLVKNKXKLTQSKIWDVVEKADIGCTPGSGDKYAGVPSDAGLTFTSSGGQTAQRA 657
Qy 659 ELQCPQPAARRRSVOLTEKMDKVGKY-PKELKCCEDGKRENMRFSCQRTFRFSLG 717
Db 658 DPECAKPAARRRSVOLTEKMDKVGKY-PKELKCCEDGKRENMRFSCQRTFRFSLG 717
Qy 718 EACKVPLDCNVITELRKHARASHGLARNDLEDIIAENIVSRSEFFESLWNVED 777
Db 718 EACKVPLDCNVITELRKHARASHGLARNDLEDIIAENIVSRSEFFESLWNVED 777
Qy 778 LKEPKNGISTKLMNIFLKDSTITTEILAVSDSKGICVADPFVTVWQDFFDILRLPY 837
Db 778 LKEPKNGISTKLMNIFLKDSTITTEILAVSDSKGICVADPFVTVWQDFFDILRLPY 837
Qy 838 SVANEQVEIRAVLYNKRQELKRVVLLHNPAPCSLATTKRHQQTITPPKSSLSVP 897
Db 838 SVANEQVEIRAVLYNKRQELKRVVLLHNPAPCSLATTKRHQQTITPPKSSLSVP 897
Qy 898 YVIVPLKTGLQEVVEAAVYHFI SDGVPKSLKYVPEGRVKNKTVAVTLDPELREGV 957
Db 898 YVIVPLKTGLQEVVEAAVYHFI SDGVPKSLKYVPEGRVKNKTVAVTLDPELREGV 957
Qy 958 QKEDIPPADLSDQVPTSETRILLQGTPEVAQVTEADVADELKHLIVTPSGCGEQNWIG 1017
Db 958 QKEDIPPADLSDQVPTSETRILLQGTPEVAQVTEADVADELKHLIVTPSGCGEQNWIG 1017
Qy 1018 MTPVIAVHYLDETEQWEGLEKQGALELKKGTQOLAFROPSSAFAPVAPSTW 1077
Db 1018 MTPVIAVHYLDETEQWEGLEKQGALELKKGTQOLAFROPSSAFAPVAPSTW 1077
Qy 1078 LTAVVVKVSLAVNLIAIDSLGCAVKVLLILEKQKPGVQFQEDAPVHQEMIGLNNN 1137
Db 1078 LTAVVVKVSLAVNLIAIDSLGCAVKVLLILEKQKPGVQFQEDAPVHQEMIGLNNN 1137
Qy 1138 EADVSLTAFVLIQALQEARLCEQVNSLPGSINKAGEYLESLNLPYVVALAGYALA 1197
Db 1138 EADVSLTAFVLIQALQEARLCEQVNSLPGSINKAGEYLESLNLPYVVALAGYALA 1197
Qy 1198 QMGLKGPLANKFLTAKQNRWEDPGKLYNVEATSVALLALQLKDFDVPVPPVRLN 1257
Db 1198 QMGLKGPLANKFLTAKQNRWEDPGKLYNVEATSVALLALQLKDFDVPVPPVRLN 1257
Qy 1258 EORYGGYGGYSTQATFWFQALAOYKQADPHOEINLVDLSQLPSRSSKITHRWESAS 1317
Db 1258 EORYGGYGGYSTQATFWFQALAOYKQADPHOEINLVDLSQLPSRSSKITHRWESAS 1317
Qy 1318 LLRSBETKEGFTVTAEGKQGTLSVVTWYHAKAKOQLTCKNFKLVTKPAPETKRP 1377
Db 1318 LLRSBETKEGFTVTAEGKQGTLSVVTWYHAKAKOQLTCKNFKLVTKPAPETKRP 1377
Qy 1378 QDAKNTMILICTRYRGDODATMSILDISMGTGAPDITDILKOLANGVDRIYSKYELDKA 1437
Db 1378 QDAKNTMILICTRYRGDODATMSILDISMGTGAPDITDILKOLANGVDRIYSKYELDKA 1437

Qy 1438 FSDNTLIIYLDKVSHEDEDDCLAFKVHOFYFVELIOPGAVKYVAYNLEESCTREYHPEK 1497
Db 1438 FSNKNTLIIYLEKISHSEEDCLSFVKHOFYFVELIOPGAVKYVAYNLEESCTREYHPEK 1497
Qy 1498 EDGKLNLCRDELCRAEENCFIQKSDDKVTLEERLDKACBPGVDYVYKTRVKVQLSND 1557
Db 1498 DCGMLSLCHNEMCRCAEENCFMHQSDQVSLNERLDKACBPGVDYVYKTRVKVQLSND 1557
Qy 1558 FDEVIMAEOTIKSGSDEVQVGOORTFISPIKREALKLEEKKHYLMGLSGDFWGEKPN 1617
Db 1558 FDEVIMAEOTIKSGSDEVQVGOORTFISPIKREALKLEEKKHYLMGLSGDFWGEKPN 1617
Qy 1618 LSYIIGKDTWVEHWPBEDECDSENOKQODLGAFTESMWVFGCPN 1663
Db 1618 TSYIIGKDTWVEHWPBEDECDSENOKQODLGAFTESMWVFGCPN 1663

RESULT 4
CO3_CAVPO
ID CO3_CAVPO STANDARD; PRT; 1666 AA.
AC P12387;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Complement C3 precursor [Contains: C3a anaphylatoxin].
GN Name=C3;
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9030798; PubMed=1973176;
RA Auerbach H.S., Burger R., Dodds A., Colten H.R.;
RT "Molecular basis of complement C3 deficiency in guinea pigs.";
RL J. Clin. Invest. 86:96-106(1990).
RN [2]
RP SEQUENCE OF 576-753.
RX MEDLINE=89113342; PubMed=3064079;
RA Gerard N.P., Lively M.O., Gerard C.;
RT "Amino acid sequence of guinea pig C3a anaphylatoxin.";
RL Biochemistry 22:942-947(1983).
RN [3]
RP SEQUENCE OF 993-1032.
RX MEDLINE=8317889; PubMed=6838833;
RA Thomas M.L., Tack B.F.;
RT "Identification and alignment of a thiol ester site in the third component of guinea pig complement.";
RL Biochemistry 22:942-947(1983).
CC -!- FUNCTION: C3 plays a central role in the activation of the complement system. Its processing by C3 convertase is the central reaction in both classical and alternative complement pathways. After activation C3b can bind covalently via its reactive thiolester to cell surface carbohydrates or immune aggregates. CC -!- FUNCTION: Derived from proteolytic degradation of complement C3, C3a anaphylatoxin is a mediator of local inflammatory process. It induces the contraction of smooth muscle, increases vascular permeability and causes histamine release from mast cells and basophilic leukocytes. CC -!- SUBUNIT: C3 precursor is first processed by the removal of 4 Arg residues, forming two chains, beta and alpha, linked by a disulfide bond. C3 convertase activates C3 by cleaving the alpha chain, releasing C3a anaphylatoxin and generating C3b (beta chain + alpha' chain). CC -!- SIMILARITY: Contains 1 anaphylatoxin-like domain. CC -!- SIMILARITY: Contains 1 NTR domain. CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. CC There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC	EMBL; M34054; AAA37038.1; -	Db	120	VLVSLQGYLFIOQDTXITVPGSTVLVRIFTVDSLLPVGRITIIIVITETPGIPIKRDTL	179
CC	DR	Qy	181	SSONQLGVLPLSDWIDIPELVNMGMQKIRAYIYENSPOQVFSTEPEVREYVLPSPFIVEPTE	240
DR	PIR; A37156; C3GP.	Db	180	SSNNQHGILPLSNWIDIPELVNMGMQKIRAYIYENSPOQVFSTEPEVREYVLPSPFIVEPTE	239
DR	HSSP; P01026; 1QOF.	Qy	241	KFYIYNEKGLVITITARELYGKKEGTAFTVIFGLDGEQRIISLPESLKRIPIEDGSEV	300
DR	InterPro; IPR002890; A2M.N.	Db	240	KFYIIDDPKGLVNIITARELYGKKEGTAFTVIFGLDGEQRIISLPESLKRIPIEDGSEV	299
DR	InterPro; IPR009048; AM_receptor_bind.	Qy	301	VLSRKVLDDGVQNPRAEDLVGKSLVYSATVILHSGSDMVQAFERSGIPITVTPYQIHFTKT	360
DR	InterPro; IPR000020; Anaphylatoxin.	Db	300	VLSRQVLLDGVQNPRAEDLVGKSLVYSATVILHSGSDMVQAFERSGIPITVTPYQIHFTKT	359
DR	InterPro; IPR009864; Invasin_intimin.	Qy	361	PKYFKGMPDLMVFTVTPDGPSPAYRVPVAVAGEVTGSLTQDGVAKLSINTHPSQPL	420
DR	InterPro; IPR001599; Macrogloblina2.	Db	360	PKYFKGMPDLMVFTVTPDGPSPAYRVPVAVAGEVTGSLTQDGVAKLSINTHPSQPL	418
DR	InterPro; IPR001134; Netrin.C.	Qy	421	SITVTRTKQELSEAEQATETMQALPYSTVGNNSNNYLHLSVLRTLRPGETLVNPLRMD	480
DR	InterPro; IPR008930; Terp_cyc_toroid.	Db	419	SITVTRTKQELSEAEQATETMQALPYSTVGNNSNNYLHLSVLRTLRPGETLVNPLRMD	478
DR	InterPro; IPR008993; TIMP_like.	Qy	481	RAHEAKIRYVYVILNNKGLLKAGQVREPQGLVPLSLITTDPIPSFRLVAYVYTLGA	540
DR	Pfam; PF00207; A2M.N.1.	Db	479	PQGEAKIRYVYVILNNKGLLKAGQVREPQGLVPLSLITTDPIPSFRLVAYVYTLGA	538
DR	Pfam; PF01835; A2M.N.1.	Qy	541	SGQREVWADSVWVVDVXDSVGSJLVKSGQSEPD-----RQPVGQOQMTLIEGDHGAVVL	595
DR	Pfam; PF01821; ANATO.1.	Db	539	SAQREVWADSVWVVDVXDSVGSJLVKSGQSEPD-----RQPVGQOQMTLIEGDHGAVVL	598
DR	Pfam; PF01753; NTR.1.	Qy	596	VAYDKGVFLVNNKNTKLSKIWDVVEKADIGCTPGSGDKYAGVFSAGLTSS--SGOOT	654
DR	PRINTS; P000004; ANAPHYLATOXIN.	Db	599	VAYDKGVFLVNNKNTKLSKIWDVVEKADIGCTPGSGDKYAGVFSAGLTSS--SGOOT	658
DR	ProDom; P003264; Anaphylatoxin.1.	Qy	655	AQRAELQCPQAPARRRRSVOLTEKMDKVGYK-PKELRKCCEDGMRNEMPFSCORTRF	713
DR	PROSITE; PS00477; ALPHA_2_MACROGLOBULIN.1.	Db	659	AQREGLDCPKPAARRRRSVOLTEKMDKVGYK-PKELRKCCEDGMRNEMPFSCORTRF	718
DR	PROSITE; PS01177; ANAPHYLATOXIN_1.1.	Qy	714	ISLGEACKVFLDCCNVITELRROHARASHLGLARSLNDEDIIEENIVSSSEPPESLW	773
DR	PROSITE; PS01178; ANAPHYLATOXIN_2.1.	Db	719	VSLGEACVAFDCCCTYMAQLRQRRRQNLGLARSLNDEDIIEENIVSSSEPPESLW	778
DR	PROSITE; PS01189; NTR.1.	Qy	774	NVEDLKEPPKNGISTKLMNIFLKDSITWEILAVSMSSDKGICVADPEVTVMDFFIDL	833
KW	Complement alternate pathway; Complement pathway;	Db	779	TIBELKEPERNGISTKLMNIFLKDSITWEILAVSMSSDKGICVADPEVTVMDFFIDL	838
KW	Direct protein sequencing; Glycoprotein; Inflammatory response;	Qy	834	RLPVSVVRNEQVEIRAVLVNVEQNOELKVRVELLNPAFCSLATTKRHHQCTIIPPKSS	893
KW	Pitama; Signal; Thioester bond.	Db	839	RLPVSVVRNEQVEIRAVLVNVEQNOELKVRVELLNPAFCSLATTKRHHQCTIIPPKSS	898
FT	SIGNAL	Qy	894	LSVYVIVFLKTLQLOVEVKAAYVHHFISDGVKRSKLVVPEGIRNMKTVAVRTLDPERLG	953
FT	CHAIN	Db	899	VAVYVIVFLKTLQLOVEVKAAYVHHFISDGVKRSKLVVPEGIRNMKTVAVRTLDPERLG	958
FT	CHAIN	Qy	954	REGVQKEDIPPADLSDQVDPDTESETRILLQGPFPVQAQMTEDAVDAERLKHLLVTPSGGEQ	1013
FT	CHAIN	Db	959	REGVQKEDIPPADLSDQVDPDTESETRILLQGPFPVQAQMTEDAVDAERLKHLLVTPSGGEQ	1018
FT	CHAIN	Qy	1014	NMIGMPTTVIAVHYLDETEQWEKFGLEKRGQALELIEKGYTQQLAFRQPSAFVAKRA	1073
FT	CHAIN	Db	1019	NMIGMPTTVIAVHYLDETEQWEKFGLEKRGQALELIEKGYTQQLAFRQPSAFVAKRA	1078
FT	CHAIN	Qy	1074	PSVTLVAVYVYVFLSALVNLIAIDSOVLCGAVKWLILEKQKPDGQVQEDAPVHDEMIGGL	1133
FT	CHAIN	Db	1079	SSVTLVAVYVYVFLSALVNLIAIDSOVLCGAVKWLILEKQKPDGQVQEDAPVHDEMIGGL	1138
FT	CHAIN	Qy	1134	RNNNEKDMALTAFLVLSIQEAKDICEEQVNSLPGSITKAGDFLBANYMNLORSYTVIAG	1193
FT	CHAIN	Db	1139	RTAQEDVSLTAFLVLSIQEAKDICEEQVNSLPGSITKAGDFLBANYMNLORSYTVIAG	1198
FT	CHAIN	Qy	1194	YALAQGRKGLPLANKFLTTAKDKRWDPPGQLNVEATSVALLALLQLKDFVPPV	1253
FT	CHAIN	Db	1199	YALLERLNGATLQKFLNAATEKRWEEARQKLSVEATSVALLALLQLKDFVPPV	1258

Query Match
Best Local Similarity 77.7%; Pred. No. 0;
Matches 1297; Conservative 180; Mismatches 182; Indels 11; Gaps 7;
1 MGPSGSPSLLLLLTHPLALGPMYSITPNILRESEETVYLAHDAQGQVPTVTVH 60
1 MGPAAGSLLLLLASVSLALGPMYSITPNILRESEETVYLAHDAQGQVPTVTVH 60
61 DFPQKLVLSKTLVTPANNMGNTFTIPANREPSEKGRKFTVVOATFGTOVKEV 120
61 DFPQKLVLSKTLVTPANNMGNTFTIPANREPSEKGRKFTVVOATFGTOVKEV 119
121 VLVSLQGYLFIOQDTXITVPGSTVLVRIFTVDSLLPVGRITIIIVITETPGIPIKRDTL 180

QY	1254	RWLNQRYYGGYGSQTQTFNVFOALQYQKDA	PHQELNLDVSIQLPSRSKTHRIHW	1313			
Db	1259	RWLNQRYYGGYGSQTQTFNVFOALQYQKDA	PHQELNLDVSIQLPSRSKTHRIHW	1318			
QY	1314	ESASLLRSEETKENEGETVTAEGKGGQGLS	VVTVMYHAKADOLTONKFDLVKTIKPAPET	1373			
Db	1319	EAGSLLRSEATKONEGFKLTAKGGQGLT	SVVAVYAKTKRKVCKNFDLAVTLKPAPDT	1378			
QY	1374	EXRPDQAKNTWLIREICTRYGDDQATMSILD	ISMTMTGPAPDPTDJDJLQKLANGVDVYISKYE	1433			
Db	1379	VKKPQKASTMILGICTRYLGDQATMSILD	ISMTGTGFPDPTDJDJLQKLANGVDVYISKYE	1438			
QY	1434	LDKASDNTLIIYLDKVKSHDEDDCLAFKVH	QVFNVELLQPGAVKYAYNLSRSTRFY	1493			
Db	1439	MNKDFS-KNTJIIYLDKVKSHSECLSFKTHQ	FNVLQPGSVKYVSYNLDSTCTQFY	1497			
QY	1494	HPEKEDGKLNKLCRDELCRCAEENCFTQKSD	DKVTLERLDRKACPGVYVYVKTLLVKVQ	1553			
Db	1498	HPEKEDGMLNKLCHKDLRCRCAEENCFTQ	LPEKTIILDERLEXACEPGVYVYVKTLLKWE	1556			
QY	1554	LSNDEDEYMAIEQTITSGSDVQVGOORFT	SPISKRCALKEKKHLMWGLSSDPNG	1613			
Db	1557	LSDDDFEYIMTIEQVYKSGSDVQVQKREF	SHIKCDRLHKLXEGKHYLMWGLSSDLNG	1616			
QY	1614	EXPNLSYIIGKDTVWEHPBEDEQDSEENKQ	QCDLGAFTESMYVFGCPN	1663			
Db	1617	ERPNNSYIIGKDTVWEAPRAECQDSEENKQ	QCDLGTENNMYVFGCPN	1666			
RESULT 5							
Q80XP1	PRELIMINARY; PRT; 1663 AA.						
AC	Q80XP1						
DT	01-JUN-2003	(TRENBLrel. 24, Created)					
DT	01-JUN-2003	(TRENBLrel. 24, Last sequence update)					
DT	01-MAR-2004	(TRENBLrel. 26, Last annotation update)					
DE	Complement component 3.						
OS	Mus musculus (Mouse).						
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.						
NCBI	TaxID=10090;						
RN	[1]_SEQUENCE FROM N.A.						
RN	STRAIN=FVB/N; TISSUE=Liver;						
RC	MEDLINE=22338857; PubMed=12477932;						
RX	Strausberg R.U., Fingold E.A., Grouse L.H., Derge J.G.,						
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,						
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,						
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,						
RA	Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,						
RA	Stapleton M., Soares M.B., Bonaldi M.P., Casavant T.L., Schetz T.E.,						
RA	Raha S.S., Locuellar N.A., Peters G.J., Abramson R.D., Mullany S.J.,						
RA	Rask S.A., Worley K.C., McKernan K.J., Malek J.A., Gumaratne P.H.,						
RA	Richardson D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,						
RA	Vallalon D.K., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,						
RA	Fahey J., Heltan E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,						
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,						
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,						
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,						
RA	Krzyszynski M.I., Skalsoda U., Smailus D.E., Schnerch A., Schein J.E.,						
RA	Jones S.J., Warra M.A.;						
RT	"Generation and initial analysis of more than 15,000 full-length human						
RT	and mouse cDNA sequences",						
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).						
RN	[2]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN=FVB/N; TISSUE=Liver;						
RA	Strausberg R.;						
RL	Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.						
DR	EMBL; BC043338; AAH43338.1; -.						
DR	HSSP; P01026; 1QQF.						

CC residues, forming two chains, beta and alpha, linked by a
 CC disulfide bond. C3 convertase activates C3 by cleaving the alpha
 CC chain, releasing C3a anaphylatoxin and generating C3b (beta chain
 CC + alpha' chain).

CC -|- ALTERNATIVE PRODUCTS:
 CC Event=Alternative initiation;
 CC Comment=2 isoforms, Long (shown here) and Short, are produced by
 CC alternative initiation;
 CC -|- PMW: C3b is rapidly split in two positions by factor I and a
 CC cofactor to form iC3b (inactivated C3b) and C3f which is released.
 CC Then iC3b is slowly cleaved (possibly by factor I) to form C3c and
 CC C3g. Other proteases produce other fragments such as C3d or C3g.
 CC -|- SIMILARITY: Contains 1 anaphylatoxin-like domain.
 CC -|- SIMILARITY: Contains 1 NTR domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: K02782; AAC42013.1; -;
 CC EMBL: J00369; AAA37336.1; -;
 CC EMBL: J00367; AAA37336.1; JOINED.
 CC EMBL: M3032; AAA37378.1; -;
 CC EMBL: Z37998; CAA86099.2; -;
 CC PIR: A92459; C3MS.
 CC HS8P: P01026; 1QOF.
 CC MG: MGI:88227; C3.
 CC GO: GO:006954; P:inflammatory response; IMP.
 CC GO: GO:0050766; P:positive regulation of phagocytosis; IMP.
 CC InterPro: IPR002890; A2M N.
 CC InterPro: IPR003048; AK_receptor_bind.
 CC InterPro: IPR000020; Anaphylatoxin.
 CC InterPro: IPR001840; Anaphylatoxin.
 CC InterPro: IPR008964; Invasin_inflam.
 CC InterPro: IPR001599; MacroglobulinA2.
 CC InterPro: IPR001134; Netrin_C.
 CC InterPro: IPR008930; Tarp_cyc_toroid.
 CC InterPro: IPR008993; TIMP_like.
 CC Pfam: PF00207; A2M; 1.
 CC Pfam: PF01835; A2M_N; 1.
 CC Pfam: PF01821; ANATO; 1.
 CC Pfam: PF01759; NTR; 1.
 CC PRINTS: PS00004; ANAPHYLATOXN.
 CC PRODOM: PD003264; Anaphylatoxin; 1.
 CC PROSITE: PS00477; ALPHA_2_MACROGLOBULIN; 1.
 CC PROSITE: PS01177; ANAPHYLATOXIN_1; 1.
 CC PROSITE: PS01178; ANAPHYLATOXIN_2; 1.
 CC PROSITE: PS01019; NTR; 1.
 CC KW Alternative initiation; Complement alternate pathway;
 KW Complement pathway; Direct protein sequencing; Glycoprotein;
 KW Inflammatory response; Plasma; Signal; Thioester bond.

FT SIGNAL 1 24
 FT CHAIN 25 1663 Complement C3, isoform Long.
 FT CHAIN 1129 1663 Complement C3, isoform Short.
 FT INIT MET 1129 1129 For isoform Short.
 FT CHAIN 25 666 Complement C3 beta chain.
 FT CHAIN 671 1663 Complement C3 alpha chain.
 FT CHAIN 671 1663 C3a anaphylatoxin.
 FT PEPTIDE 749 1663 Complement C3b alpha' chain.
 FT CHAIN 749 954 Complement C3c fragment.
 FT CHAIN 955 1303 Complement C3dg fragment.
 FT CHAIN 955 1001 Complement C3g fragment.
 FT CHAIN 1002 1303 Complement C3d fragment.
 FT PEPTIDE 1304 1320 C3f fragment.
 FT DOMAIN 693 728 Anaphylatoxin-like.
 FT SITE 1518 1661 NTR.
 FT SITE 748 749 Cleavage (by C3 convertase).
 FT SITE 1303 1304 Cleavage (by factor I).
 FT SITE 1320 1321 Cleavage (by factor I).

FT DISULFID 559 816 Interchain (By similarity).
 FT DISULFID 626 661 By similarity.
 FT DISULFID 693 720 By similarity.
 FT DISULFID 694 727 By similarity.
 FT DISULFID 707 728 By similarity.
 FT DISULFID 973 1513 By similarity.
 FT DISULFID 1101 1158 By similarity.
 FT DISULFID 1389 1489 By similarity.
 FT DISULFID 1389 1458 By similarity.
 FT DISULFID 1506 1511 By similarity.
 FT DISULFID 1518 1590 By similarity.
 FT DISULFID 1537 1661 By similarity.
 FT DISULFID 1637 1646 By similarity.
 FT CARBOHYD 939 939 N-linked (GlcNAc...).
 FT CARBOHYD 1617 1617 N-linked (GlcNAc...).
 FT CROSSLINK 1010 1013 Iso-glutamyl cysteine thioester (Cys-Gln)
 FT CROSSLINK 1010 1013 (By similarity).
 FT SEQUENCE 1663 AA; 166482 MW; D55546CC769BEA19 CRC64;
 Query Match 78.4%; Score 6749; DB 1; Length 1663;
 Best Local Similarity 77.1%; Pred. No. 0;
 Matches 1284; Conservative 179; Mismatches 197; Indels 6; Gaps 5;

QY 1 MGPTSGPSL--LILLTHLPLALGSPWYSIITPNILRLSEETMVLAEADAGDGVFVTVT 58
 DB 1 MGPSGSQLVLLVLLASSPLALGIPWYSIITPNVLRLESETIVLEADAGDQDIPVTVT 60
 QY 59 VDFPFGKVLJLSSEKTVLTPTAHMGNVTFTPANREFKSEKGRNKFVTVQATFGTQVE 118
 DB 61 VQDFL-KRQVLTSEKTVLTGASGHLRSVSIKIPASKEFNSDKGHHKYVTVVANFGETVVE 119
 QY 119 KWLVLISQSGYLFQTDTKTYTPGSTVLYRIFVFNHKLIPVGRVWNIENPEGIPVKQD 178
 DB 120 KAVWVFQSGYLFQTDTKTYTPGSTVLYRIFVFNHKLIPVGRVWNIENPEGIPVKRD 179
 QY 179 SLSSQNLGVLPISWDIPELVNNGWKIRAYENSQQVSTEFVEYKVIYVPSFEVIVEP 238
 DB 180 ILSSNQHGLFUSWNPILVNNQGWKIRAFYEHAPKQIFSAEFVEYKVIYVPSFEVIVEP 239
 QY 239 TEKPYIYNEKGLVITITAFLYGKKEVGTAFVFIQDGEQRIQLPSLKRKRIPIEDSG 298
 DB 240 TETFYIDDPNGLEVSIIAKFLYKKNVDGTAFVFIQDGEQRIQLPSLKRKRIPIEDSG 299
 QY 299 EWLRSKVLDDGQNFPAEDLVGKSLYVSATVLLHSGSDMVQAESG:PIVTSVQIHT 358
 DB 300 DAVLRKVLMEGVAPSNAADLVGKSLYVSATVLLHSGSDMVQAESG:PIVTSVQIHT 359
 QY 359 KTPKYFKGMPFDLMVFTVNPDGSPAYRVFVAVQGEDTVQSLTQDGVAKLSINTHPSQK 418
 DB 360 KTPKFFKAMPFDLMVFTVNPDGSPASKVLVTVTQSGN-AKALTQDQGVAKLSINTHPSRQ 418
 QY 419 PLSTVTKKQELSEAEQARTVQALPYSTVGNLHSLVRLTELPGETLVNVELLR 478
 DB 419 PLTITVTKQDTPESQATKWEAHYPTSMNSNNYLHLSRMLKPGDNLNWNFLR 478
 QY 479 MDRAHEAKIRYYTYLIMNKGKLLKAGQVREPGQDVLWLPISITTDIFPSFRLVAYTYLI 538
 DB 479 TDPGHEAKIRYYTYLIMNKGKLLKAGQVREPGQDVLWLPISITTDIFPSFRLVAYTYLI 538
 QY 539 GASGQREWADSVWVDVSKVSLVYVQSGSDRQVPGQQQMTLKIEGDHARVWLVAV 598
 DB 539 GASGQREWADSVWVDVSKVSLVYVQSGSDRQVPGQQQMTLKIEGDHARVWLVAV 597
 QY 599 DKGVFLVNNKNTQSKINVDVEKADIGCTPGSGKDYAGVPSDAGLTTSSSGQQTQARA 658
 DB 598 DKGVFLVNNKNTQSKINVDVEKADIGCTPGSGKDYAGVPSDAGLTTSSSGQQTQARA 657
 QY 659 ELQCPQFAARRRSVOLTEKRMVKGY-PKELRKCCEDGMEENPMWRSFCORRTFISLG 717
 DB 658 DLECTKPAARRRSVOLMERMDKAGQYTDKGLRKCCEDGMDIPMYSCORRRLITQ 717
 QY 718 EACKVFLDCNYYTELARQHAASHGLARSNLEDDIIEENIVSRSEFFESMIMWYED 777

Db 718 ENCIAKAFIDCCWHITKURBOHRDHVUGLARSLEBEDIIPBEDISRSHPFQSWLWITIE 777
QY 778 LKEPPKNGISGLKMNIFLKDITWEILAVSMGDKGICVADPFVTVNQDFFIDRLRPY 837
Db 778 LKEPPKNGISGLKMNIFLKDITWEILAVSMGDKGICVADPFVTVNQDFFIDRLRPY 837
QY 838 SVNRNEQVEIRAVLYNRQCEKLVKRVVLELHNFACSLATTKERHOOTITIPKSSLSVP 897
Db 838 SVNRNEQVEIRAVLYNRQCEKLVKRVVLELHNFACSLATTKERHOOTITIPKSSLSVP 897
QY 898 YVIVPLKGTQVEVKAAYVHFTSDGVKSLKVPKGVKIMKNTAVRTLDPERLREGV 957
Db 898 YVIVPLKGTQVEVKAAYVHFTSDGVKSLKVPKGVKIMKNTAVRTLDPERLREGV 957
QY 958 QKEDIPADLSDQVDPDTESTRILLQGTPTVAQMTEDAVDAERLKHILVTPSCGQGNMIG 1017
Db 958 QKEDIPADLSDQVDPDTESTRILLQGTPTVAQMTEDAVDAERLKHILVTPSCGQGNMIG 1017
QY 1018 MPTVIAVHYLDETEQWEKFGLEKROGALBLIKKGYTQQLAFQPSAFAAFVKAAPSTW 1077
Db 1018 MPTVIAVHYLDETEQWEKFGLEKROGALBLIKKGYTQQLAFQPSAFAAFVKAAPSTW 1077
QY 1078 LTAYVVKVFLAVNLIAIDSVOLCGAVKWLILEKQKPDGVFOBDAPVHOEMIGLRNN 1137
Db 1078 LTAYVVKVFLAVNLIAIDSVOLCGAVKWLILEKQKPDGVFOBDAPVHOEMIGLRNN 1137
QY 1138 EKDWALTAFLVLSLOEAKDICEQVNSLPGSITKAGDFLEANNYMLQRYTVAIAGYALA 1197
Db 1138 EADVSLTAFLVLSLOEAKDICEQVNSLPGSITKAGDFLEANNYMLQRYTVAIAGYALA 1197
QY 1198 QMGLKPLNKLFTTAKDKNRVEDPGKQLYNVEATSYALLALLQKDFDPPVVRWLN 1257
Db 1198 QMGLKPLNKLFTTAKDKNRVEDPGKQLYNVEATSYALLALLQKDFDPPVVRWLN 1257
QY 1258 BORYGGYGSTQATFWFOALQYQKADPHOELNLDVSLQIPRSSKITHRIHWSAS 1317
Db 1258 BORYGGYGSTQATFWFOALQYQKADPHOELNLDVSLQIPRSSKITHRIHWSAS 1317
QY 1318 LLRSEETKQNEAFSLTAKGKRGTLVSVVYHAKLSKVTCKKFLRVSVIRPAPITAKKP 1377
Db 1318 LLRSEETKQNEAFSLTAKGKRGTLVSVVYHAKLSKVTCKKFLRVSVIRPAPITAKKP 1377
QY 1378 QDAKNTMLEICTHYRGDQATWSILDSMTGAPDPTDQLQANGVDVRYISKYELDKA 1437
Db 1378 QDAKNTMLEICTHYRGDQATWSILDSMTGAPDPTDQLQANGVDVRYISKYELDKA 1437
QY 1438 FSNKNTLIYLEKISHTEDCLTFKHQYFNVGLIQGSKVYVYNNLEESCTRFYHPEK 1497
Db 1438 FSNKNTLIYLEKISHTEDCLTFKHQYFNVGLIQGSKVYVYNNLEESCTRFYHPEK 1497
QY 1498 EDGKLNLCRDELRCRAENCFTQKSDDKVTLLEERLDKACQGVYVYKTRLVKVLQND 1557
Db 1498 EDGKLNLCRDELRCRAENCFTQKSDDKVTLLEERLDKACQGVYVYKTRLVKVLQND 1557
QY 1558 FDEYIMAEQTIKSGSDEVQVQORTFISPIKREALKLEKXKHYLMWGLSDFWGEKPN 1617
Db 1558 FDEYIMAEQTIKSGSDEVQVQORTFISPIKREALKLEKXKHYLMWGLSDFWGEKPN 1617
QY 1618 LSYIIGKDTVWEHHPEDQEDQENQKQCODLCAFTSMVVGCPN 1663
Db 1618 LSYIIGKDTVWEHHPEDQEDQENQKQCODLCAFTSMVVGCPN 1663

RESULT 7

Q9GKPL PRELIMINARY; PRT; 1661 AA.
AC Q9GKPL
DT 01-MAR-2001 (TreeBrel. 16, Created)
DT 05-JUL-2004 (TreeBrel. 16, Last sequence update)
DE Complement component C3 (Complement C3)
OC Sus scrofa (Pig)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=21313047; PubMed=11419349;
RA Wimmers K., Kechay S., Ponsuksili S., Hardge T., Yerle M.,
Schellander K.;
RT "Polymorphic sites in exon 15 and 30 of the porcine C3 gene."
RL Anim. Genet. 32:46-47(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=22444329; PubMed=13557058;
RA Wimmers K., Ponsuksili S., Schellander K., Ponsuksili S.;
RT "Molecular characterization of the pig C3 gene and its association
with complement activity."
RL Immunogenetics 54:714-724(2003).
DR EMBL; AF154933; AAG40565.1; -;
DR EMBL; AJ494748; CAD38823.2; -;
DR HSP; P01026; IQQF.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004866; F:endorpeptidase inhibitor activity; IEA.
DR GO; GO:0005956; P:complement activation; IEA.
DR GO; GO:0006954; P:inflammatory response; IEA.
DR InterPro; IPR002890; A2M N.
DR InterPro; IPR009048; AM_Receptor_bind.
DR InterPro; IPR000020; Anaphylatoxin.
DR InterPro; IPR001840; Anaphylatoxin.
DR InterPro; IPR008964; Invasin_intimin.
DR InterPro; IPR001599; MacroglobinA2.
DR InterPro; IPR001134; Netrin C.
DR InterPro; IPR008930; Texp_cyc_toroid.
DR InterPro; IPR008993; TIMP_like.
DR Pfam; PF00207; A2M; 1.
DR Pfam; PF01835; A2M N; 1.
DR Pfam; PF01821; ANATO; 1.
DR Pfam; PF01759; NTR; 1.
DR PRINTS; PRO0004; ANAPHYLATOXN.
DR ProDom; PD003264; Anaphylatoxin; 1.
DR SMART; SM00104; ANATO; 1.
DR SMART; SM00643; C345C; 1.
DR PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
DR PROSITE; PS00189; NTR; 1.
SQ SEQUENCE 1661 AA; 186805 MW; 4899D0914BE3310C CRC64;

Query Match

78.1%; Score 6724; DB 2; Length 1661;

Best Local Similarity 76.5%; Pred. No. 0;

Matches 1272; Conservative 189; Mismatches 200; Indels 2; Gaps 2;

QY 1 MGSTSPSLLLLHLLPLALGSPMYSLITNLRLESEETMVLAEHDAQGDVPTVTVH 60

Db 1 MGSTSPRLLLLLLSLPLALGDPPIYITITPNVLRLESEEMVLEAHEGGDIRSVTVH 60

QY 61 DFFGKKLVLSSEKTVLTPTNMGVNTFTIPANREFKSEKGRKFTVTVQATFGTVVEKV 120

Db 61 DFFAKQVLSSETTTNNANNVLTSTNLIKIPASKEFKSEKQ-HKFTVTVQALFGVQVEKV 119

QY 121 VLVSLSQGLVLTQDKTIYTSGLVRLFTVNHKLPLVGTVMVNTENEGIPVKODSL 180

Db 120 VLVSLSQGLVLTQDKTIYTSGLVRLFTVNHKLPLVGTVMVNTENEGIPVKODSL 179

QY 181 SSQNLGVLPLSWDIPELVNMGMKIRAYENSPQVSTFEFVKEVYLPFSFVIVEPTE 240

Db 180 SSNQPGILALSWNIPELVNMGMKIRAHYEDAPQVFSAEFEVKEVYLPFSFVIVEPSE 239

QY 241 KFYIYNEKLEVTITAREFLYKKGVEGTAFVIFGIQDGEORISLIPESLKRIPIEDSGSEV 300
DB 240 KFYIYDIDPGLTUNIIAREFLYGESVDGTAFVIFGVQDQORISLSQSLTRVPIIDGTGEA 299
QY 301 VLSKVLGQVQNPRAEDLVGSLVYSATVILHSGSDMVQASRGIPVTSPIYHFTKT 360
DB 300 TLSGVLLNGVHSSVNDLVGSIYVSIVTILNSGSDMVEARTGIPVTSPIYHFTKT 359
QY 361 PKYFKGMPDLVVFVNTDPSPAYRVVAVGEBTVQSLTQGDGVAKLSINTHPSQPL 420
DB 360 PKYFKGMPDLVVFVNTDPSPAYRVVAVGEBTVQSLTQGDGVAKLSINTHPSQPL 418
QY 421 SITVTRKQELSAEQATRMQALPYSTVGNENMYLHLSVLRTLPGETLNVNFLRMD 480
DB 419 PITVTRKQELSAEQATRMQALPYSTVGNENMYLHLSVLRTLPGETLNVNFLRMD 478
QY 481 RAHEAKRYTYTILMKGRLKAGQVREPQGLVVLPLSTITDFIPSRVLVAYTIGA 540
DB 479 PGYQDKIRYTYTILMKGRLKAGQVREPQGLVVLPLSTITDFIPSRVLVAYTIGA 538
QY 541 SGQREVADSVWVVDKSDSCVGLSVKSGQSDQVPGQOMTLKTEGHDGARVVLVAVDK 600
DB 539 NGQREVADSVWVVDKSDSCVGLSVKSGQSDQVPGQOMTLKTEGHDGARVVLVAVDK 598
QY 601 GVFVNLKKNLTKOSKWDVVEKADIGCTPGSGKDVAGVPSDAGLTFITSSSQOQTAQAE 660
DB 599 GVFVNLKKNLTKOSKWDVVEKADIGCTPGSGKDVAGVPSDAGLTFITSSSQOQTAQAE 658
QY 661 QCPQFAARRRSVQLTEKMDKVGKPKELRCCEDGRENPNRPFSCORRTFTSLGEAC 720
DB 659 ECPKFAARRRSVQLTEKMDKVGKPKELRCCEDGRENPNRPFSCORRTFTSLGEAC 718
QY 721 KVFLLDCNVITLRRQHARASHGLGARSNLEDIIARENVISSEPPESWLVNVELKE 780
DB 719 VKAPLDCCEVIAKLQCHGRNPLGLARSNLEDIIARENVISSEPPESWLVNVELKE 778
QY 781 PPKNGISTKLMNIFLKDSITTWELLAVSMDKKGICVADPFVTVMOQDFIDLELPYSV 840
DB 779 PPKNGISTKLMNIFLKDSITTWELLAVSMDKKGICVADPFVTVMOQDFIDLELPYSV 838
QY 841 RNEQVEIRAVLYNQVQELKVELLHNPACSLATTKRHOOTITIPKSSVPEYVI 900
DB 839 RNEQVEIRAVLYNQVQELKVELLHNPACSLATTKRHOOTITIPKSSVPEYVI 898
QY 901 VPLKTGLQEVVEKAAVYHPIFSDGVKSLKVPYEGIRMKTVAVTLPDLREGVQKE 960
DB 899 VPLKTGLQEVVEKAAVYHPIFSDGVKSLKVPYEGIRMKTVAVTLPDLREGVQKE 958
QY 961 DIPADLSQVPTSETRILQGTPEVAQWTEDAVDBELKHLIVTPSGGGEQNNIGMT 1020
DB 959 DIPADLSQVPTSETRILQGTPEVAQWTEDAVDBELKHLIVTPSGGGEQNNIGMT 1018
QY 1021 TVIAVHYLDTEQWKEFKLEKQALELKKGYTQCLAPRQPSGAFAPVYKRAPSTWLT 1080
DB 1019 TVIAVHYLDTEQWKEFKLEKQALELKKGYTQCLAPRQPSGAFAPVYKRAPSTWLT 1078
QY 1081 YVVKVFLAVNLIAIDSVLCGAVKWLILEKQKPGVFCEDAPVTHQEMIGLNNNEKD 1140
DB 1079 YVVKVFLAVNLIAIDSVLCGAVKWLILEKQKPGVFCEDAPVTHQEMIGLNNNEKD 1138
QY 1141 MALTAFLVLSQBAKIDCEQVNSLPGSTKAGDFLEANYNQLORSYTVATAGVALAOM 1200
DB 1139 VSLTAFVLLAQBAKIDCEQVNSLPGSTKAGDFLEANYNQLORSYTVATAGVALAOM 1198
QY 1201 RLKGLFLNKLFTAKDNKWEOPGKOLVNVETATSVALLALQLKDFVFPVVRWLNQ 1260
DB 1199 KLDEFFLNLKLSLAKERNWEOPGKOLVNVETATSVALLALQLKDFVFPVVRWLNQ 1258
QY 1261 YGGGYSQSTQATPVVQALAOQKADAPHOELNLDVLSQLPSRSKITHRHINESALL 1320
DB 1259 YGGGYSQSTQATPVVQALAOQKADAPHOELNLDVLSQLPSRSKITHRHINESALL 1318
QY 1321 SEETKENEGFTVTAEGKGQGTLSVVTMTHAKAKDQLTCKNFKDLKVTIKPAPETEKRPQDA 1380

DB 1319 SEETKENEGFTVTAEGKGQGTLSVVTMTHAKAKGTCKKFKLKVSIHPAPEPVKKQPEA 1378
QY 1381 KNTMILICITRYRGDOATMSILDIISMTGFPADPTDOLKOLANGVDRVYSKYELDKAFSD 1440
DB 1379 KSSNVLDICITRYLGNQDATMSILDIISMTGFPDIEDLKLSTGVDRVYSKYELNKLNS 1438
QY 1441 RNTUILLVKSSEHDDCLAFKQVQFNVVELIQPGAVKVYAYNLEESCTRFYHPEKEDG 1500
DB 1439 KNTLIIYLDKISHTLEDICISFKVQFNVVGLIQPGSVKYYSYNLDESCTRFYHPEKEDG 1498
QY 1501 KLNKLCEDELRCRAEENCFQKSDDKVLEERLDKACEPGVDYVYKTLVKVQLNSDDE 1560
DB 1499 MLNKLCHKWCRCAEENCFMWHDEEVTLLDRLERACEPGVDYVYKTLVKVQLNSDDE 1558
QY 1561 YMAIEGTIISGSDDEVQVGOQRTFISPIKCRALKLBEKKHYLMGLSSDFWGEKPNLSY 1620
DB 1559 YIMVIEGTIISGSDDEVQVGOQRTFISPIKCRALKLBEKKHYLMGLSSDFWGEKPNLSY 1618
QY 1621 IIGKDTVVEHPBEDECEQDEENKQCDLGAFTESVVFVGCEN 1663
DB 1619 IIGKDTVVEHPBEDECEQDEENKQCDLGAFTESVVFVGCEN 1661

RESULT 8
Q90633
ID Q90633 PRELIMINARY; PRT: 1652 AA.
AC Q90633;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Complement C3 precursor.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
NCBI_TaxID=9031;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEOLINE=95173423; PubMed=7532662;
RA Mavroidis M., Sunyer J.O., Lambiris J.D.;
RT "Isolation, primary structure, and evolution of the third component of
RT chicken complement and evidence for a new member of the alpha 2-
RT macroglobulin family.",
RL J. Immunol. 154:2164-2174 (1995).
DR EMBL; U16848; AAA64694.1; -
DR PIR; I50711; I50711.
DR HSP; P01024; LGHO.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004866; F:endorpeptidase inhibitor activity; IEA.
DR GO; GO:0006956; P:complement activation; IEA.
DR GO; GO:0006954; P:inflammatory response; IEA.
DR InterPro; IPR002890; A2M.N.
DR InterPro; IPR009048; A2M receptor bind.
DR InterPro; IPR000020; Anaphylatoxin.
DR InterPro; IPR001840; Anaphylatoxin.
DR InterPro; IPR001599; MacroglobulinA2.
DR InterPro; IPR001134; Netrin_C_toroid.
DR InterPro; IPR008930; Terp_cyc_toroid.
DR InterPro; IPR008993; TIMP_like.
DR Pfam; PF00207; A2M; 1.
DR Pfam; PF01835; A2M.N; 1.
DR Pfam; PF01821; ANATO; 1.
DR Pfam; PF01759; NTR; 1.
DR PRINTS; PR00004; ANAPHYLATOXIN.
DR ProDom; PD003264; Anaphylatoxin; 1.
DR SMART; SM00104; ANATO; 1.
DR SMART; SM00643; C345C; 1.
DR PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR PROSITE; PS01177; ANAPHYLATOXIN_1; UNKNOWN_1.
DR PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
DR PROSITE; PS01189; NTR; 1.

KW Signal.
 FT CHAIN 1 16 Potential.
 FT CHAIN 644 1652 Complement C3 alpha chain.
 FT CHAIN 17 639 Complement C3 beta chain.
 SQ SEQUENCE 1652 AA; 184085 MW; 9020C49FE127BD1F CRC64;

Query Match
 Best Local Similarity 52.5%; Score 4518.5; DB 2; Length 1652;
 Best Local Similarity 53.4%; Pred. No. 2.1e-260;
 Matches 896; Conservative 305; Mismatches 435; Indels 41; Gaps 24;

QY 1 MGPTSGPGLLLLLTHPLALGSP--MYSIIITNPLRLESEETMVLAEADAQDVPVTV 58
 DB 1 MGLLLPLLLGLVLLH--AVPTPAQVMTVTPAVLRLDTDEKVLDEAPGLSAPTEANL 57
 QY 59 VHDFFGKLVLSSEKTVLTPATNMGVNTFTIPANREPKSEKGRNKFVTVQATFGQVVE 118
 DB 58 VQDFPQKRVLFQVKQLNPAEGWMAIATKVPV-KLLPPVKGK-FVSVARVQGVTL 115
 QY 119 KVLVLSQSGVLFOTQCTKTVTPGSTVLYRIFTVNHKLLPVGTVMVNIENEGIPKOD 178
 DB 116 KVLVLSQSGHIFQTDKPIVTPGSTVLSRLFALSHFMQLLTVIVEKTPDNVIRKQV 175
 QY 179 SLSSQNLGVPLSDIPELVYMGQWKIRAYENSPQVFEFTEFEVKEYVLPSEFVTLDP 238
 DB 176 PVSSPMRNGIFSNHNLPEVSLGTWITAKFDSQDQVFEFTEFEVKEYVLPSEFVTLDP 235
 QY 239 TEKFYIYNKGLVITITARLYGKVEGTAFVIFGIQDGEORISLPESKRIPIEDGSG 298
 DB 236 QKFLYIDPDEFRVITARLYGKNLQGTAFVLFGVVDEKKTIPQSLQKRVKVDGQ 295
 QY 299 EVLSRKYVLDGVQNPRAEDLVGKSLVYSATVILHSGSDMVQAESGIPVTSPOIHT 358
 DB 296 QAVLPWMLRPPAN--LQELVGHSLVTVTVLTESGSDMVQAESGIRIVTSPIYIHT 353
 QY 359 KTEPKFKGMPDPLMVFVTPNPGSPAYR-VPVAVQGEDTVQSL--TOGQVAKLSINTHP 415
 DB 354 HTPKFKGMPDPVTVTVNEDNSPAAAGIPVKA--DNFQGLVSCRDGTAKLVNMPA 410
 QY 416 SKPLSIVTRTKQELSAEQAATMOALPYSTVGNNSNYLHLSVRLTELPGETLVNPF 475
 DB 411 KNNSVPIIVTRDQDLPPERQASQIVAEAYQSGNSGNYLHLAVASQVQPDNLPINF 470
 QY 476 LURMDRAHEAK-IRYTYLIMNKGRLKAGQVREPQDLVPLSLSTDTFIPSPRLVAY 534
 DB 471 HLKSNRDDVRKSVSYFTVLLSKGHIHVGRQREGQDQSLVTSVPTANLPSFRVAY 530
 QY 535 YTLIGASQREVAVSDVWVYKDCVGLSVYK-SQGEDRQPVPGQOMTLKEGDHGRV 593
 DB 531 YHV----RPGELIADSVWVDKTCMGLSVVRGASEADNRVHPPTMELHIEGDHKA 586
 QY 594 VLVAVDKGVFLNKKNTQSKIWDVVEKADIGCTPGSKDYAGVFSADGLFTSSGQQ 653
 DB 587 GLVAVDKAVYVLN-KNKLTSQKWDVTENSIGCTPGSGRNVGVFPADAGLSITSNVNIN 645
 QY 654 TAQRAELQCPQAPARRRSVQLTEKMDKVKY-PKELRKCCDGHRENPMRSCORRTR 712
 DB 646 TEQSEVQCAKPAKGRKSRVRLIKHGTQMAEYSDKNLRKCCDGHRENLMGYSCBKAT 705
 QY 713 FTSLSGACKVFLDCNVTILRQHARASHLGLARSNLDEDIIASENIVSRSEPPESML 772
 DB 706 VYLDKASTCEAPLSCCLYKIGIRDEE-RELQVELARSEVDDAFLSDSDITSRLFSFSL 764
 QY 773 WNVDELKBPB-KNGISTKLMIIFLKDSITTWELIAMSDDKIGICVADPEFVTVMQDFI 831
 DB 765 WQVEBELTPPNEQGSMTLPIYLKDSITTWELIAMSISSENKGLCVADPEFVTWMEFFI 824
 QY 832 DLRLPYSVVRNEQVIRAVLYNTRQNSLKVVRVLELHNPAPFCSLATTKRRHQOITIPPK 891
 DB 825 DLRLPYSAVRNEQVIRAILNYWTN-KIKVRVELMWNPALCSASTTKTRYQYIQLLEBQ 883
 QY 892 SLSVPEYIVPLKTLQSGVEVKAAYVHHFISDGVKSLKVVPGIRKWNKTAVRTDPER 951
 DB 884 SSDAVFVTVPLVLELQGHQVRAAVWNSFVSDGVKKKRLRVVPSGMRLEKTVKVELDPKT 943

QY 952 LGRGQVOKEDIPADLSDOVPDPTSETRILLQCTPVAQWTEDAVDAERLKHILVTPSGCG 1011
 DB 944 LGNNGQVEVKVKAAMLSDIVPNTSETKVSVQGNPVSILVEKATDGTGKHLIVTPSGCG 1003
 QY 1012 EQNMIGMTPTVIAVHYLOBTEQWEXFGLERQGALELILKKGYTQQLAPROPSSFAAFVK 1071
 DB 1004 EQNMIGMTPTVIAVHYLOBTEQWEXFGLERQGALELILKKGYTQQLAPROPSSFAAFVK 1063
 QY 1072 RAPSTWLTAVVVKVFLSALNLAIDSOVLCGAVKWLILEKOKPGVQEDAPVTHQBMIG 1131
 DB 1064 RPSTWLTAVVVKVFLSALNLAIDSOVLCGAVKWLILEKOKPGVQEDAPVTHQBMIG 1123
 QY 1132 GLANNEKDMALTAFLVLSLOPAKDI CBEQVNSLPGSITKAGDGFLEANYMLQSRVYAI 1191
 DB 1124 GY-HGAPSVSLTAFVLSALQESQKIKYVKSIDGSIKASDYLSRKYQSLRPTVVAL 1182
 QY 1192 AGVALAQMRKLG-PLLNKFLTTAKNURWEDPGKQLNVEATSVALLALLQLKDFDVP 1250
 DB 1183 TSVALALTGKLSEKVMKF---SKDGTWAEARNATVNIESTYVALVALLQMEKSLTG 1239
 QY 1251 FVYRWLNBOYRGVGGYSTQATFMVFOALAQYQKADPDHQLNLDVSLQSPRSKITHR 1310
 DB 1240 FVYRWLAQNYFGGYGSTQATILFVFOALAQYHVALPHEVNLNLDVSVLLPRANATYR 1299
 QY 1311 IHVESASLLRSBETKENEGFTVTAEGKQGTLSVVVMTYHAKAD-QLTCNKFDLKV 1366
 DB 1300 IENNALVARSATKLNEDFTVKAEGTGTGIMTVVTVYKAVPEKENKCNDFLRVSVED 1359
 QY 1367 IKAPATEKRPQADAKVMTLEICTRYRGDQDATWSILDISMGTGAPDPTDILKOLANGVD 1426
 DB 1360 VKAGREVEGVIRSVKIT----ICTRFLDVTDATWSILDISMGTGAPDPTDILKOLANGVD 1415
 QY 1427 RYLSKVELKAFSDRNTLIYLDKVSHEDEDCIAPKVKHQYFNVELIQPGAVKVVAYNLE 1486
 DB 1416 RYLSKEIDHALSNRNLIIYLDKVSHEDEDCIAPKVKHQYFNVELIQPGAVKVVAYNLE 1475
 QY 1487 ESCTRYHYHEKEDGKLNKLCRDELRCABENCFIQ-KSDDKVTLERLDKACEPQVDVY 1545
 DB 1476 DRCTRYHYHDPKAGQURKICHGEVC-CABENCFIRKVKNDPITVNERIDLACKPGVDVY 1534
 QY 1546 KTRLVKQVQSNDFEYMAIEQTIKSGDVEVQVQQTFFISPIKREALKLEKKHYLMW 1605
 DB 1535 KVKVATEETEPHDNYMSILTVIKVGTDENPGSNRIFVSHKQCRDALSLOKQDYLW 1594
 QY 1606 GLASDWFGEKPNLSYIIGKQVWEHVEHVEDEDEQENKQCDLGAFTESMVVSGCP 1662
 DB 1595 GLASDLWVTGSRFSYLSKDTWLEAPLUESQDADLOPLCQDFTFESDNJLVFGCP 1651

RESULT 9
 CO3_NAJNA
 ID CO3_NAJNA STANDARD; PRT; 1651 AA.
 AC Q01833;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Complement C3 precursor [Contains: C3a anaphylatoxin].
 GN Name=C3;
 OS Naja naja (Indian cobra).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroides;
 CC Elapidae; Elapinae; Naja.
 OX NCBI_Taxid=35670;
 [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Liver;
 RC MEDLINE=3056528; PubMed=1431125;
 RA Fritzinger D.C., Connolly M., Petrella E.C., Bredehorst R.,
 RA Vogel C.W.;
 RT "Primary structure of cobra complement component C3";
 RL J. Immunol. 149:3554-3562(1992).
 CC -!- FUNCTION: C3 plays a central role in the activation of the


```

Db 1010 SVIATYLDATQWENLGVDRTERAIKQIMTGVAQWVYKKAHSDHVAATNRASSSWLTA 1069
Qy 1081 YVVKVFSIANDLIA-IDQVLCGAVKMLILEKQKPGDGVFOEDAPVHQMIGELNNNEK 1139
Db 1070 YVVKVLANASNNKXDSHIEHICGGVKKWLIINROQDQGVFXENAPVHGMELGGTK-GRAP 1128
Qy 1140 DMALTAFLVLSIQEAKDICEBQVNSLPGSITKAGDFLEANNMQLQSYTVAJAGVALAQM 1199
Db 1129 BASLTAFTVLTALLESRSVCKEQINILDOSSINKATDYLKKYKQLQPYTALTAYALAA 1188
Qy 1200 GELKPLINKLTTAKDKNRHDEKQLYNVVEATSTALLALLQLKDFPVPVWRWLNQ 1259
Db 1189 DELNDRV--LMAASTGGRNREWEYNATHNIEGTSVALLALLKMKFAEVGVFWRLIDQ 1246
Qy 1260 RYGGGYSQTATFWFOALAQYQKADPDHQLNLDVSLQLPERSKITHRIHWESASLL 1319
Db 1247 KYGGTYGQTQATVWVFCALAEYEIOMPTHODLNLDSIKLPKPEVPERYSINDRNVQA 1306
Qy 1320 SEETKENGEGTVTAEGKGQGLSVVTVMYHAKK-DOLTCNKFDLKVTKPAPETEKRPQ 1378
Db 1307 RVETKLNEDTVSASGGKATMTLTVYNAQREDANVCNRPFLDVSVENVELNLKQAK 1366
Qy 1379 DAKNTMILEICTRYRGQDQATMSILDISMMTGAPDPTDDLKQANGVDRYISKYELDKAF 1438
Db 1367 GGAALRLKICTRYLGEVDSTMTIID:SMLTGFPDAEDLKLNSGVDRYISKFEIDNM 1426
Qy 1439 SDNNTLIYLDKSHSEDDCLAPKVOYENVLIQGVAVYAYXNLBESCTRFVHPKE 1498
Db 1427 AQKGVTVIYLDKSHSEDDCLAPKVOYENVLIQGVAVYAYXNLBESCTRFVHPKE 1486
Qy 1499 POKNLKCRDLRCABENCFIQSDDKVTLEBLDKACBPVGVYVYKTLVYKQVLSNDF 1558
Db 1487 TGLLNKICHGNICRABETCSLLNQKKIDQLRIQKACQNVYVYKTLRISEKDN 1546
Qy 1559 DEYMAISQTIKSGSDEVQVQOQRTFSPIKCEALKLEKKHVMWGLSSDFWGEKPNL 1618
Db 1547 DIYFMDVLEVTKGTDNRNAQAKARQYVSQRKCEALNLKLDNLYLWGLSSDLNPKWDI 1606
Qy 1619 SYIICKTQWVHNWPEDEECQDENQKQCDLGAFTESMVVFGCP 1662
Db 1607 SYLITKNTWIERWNEDEECQDENQKQCDLGAFTESMVVFGCP 1650

RESULT 10
Q91132 PRELIMINARY; PRT; 1642 AA.
AC Q91132;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cobra venom factor precursor.
OS Naja kaouthia (Monocled cobra) (Naja siamensis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Naja.
OX NCBI_Taxid=8649;
RN (1)
RN SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RX MEDLINE=9510841; PubMed=7809120;
RA Fritzinger D.C., Bredehorst R., Vogel C.W.;
RT "Molecular cloning and derived primary structure of cobra venom
factor.";
RT factor.";
RJ Proc. Natl. Acad. Sci. U.S.A. 91:12775-12779(1994).
RN (2)
RN SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RA Fritzinger D.C.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; U09669; A2468989.1; -.
DR HSP; P01026; 1QZF.
CO GO:0005576; C:extracellular; IEA.

```

```

GO:0004866; F:endorpeptidase inhibitor activity; IEA.
GO:0006956; P:complement activation; IEA.
GO:0006954; P:inflammatory response; IEA.
DR InterPro; IPR002890; A2M N.
DR InterPro; IPR009048; AM_receptor_bind.
DR InterPro; IPR000020; Anaphylatoxin.
DR InterPro; IPR001840; Anaphylatoxin.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR008964; Invasin_intimin.
DR InterPro; IPR001599; MacroglobinA2.
DR InterPro; IPR001134; Netrin_C.
DR InterPro; IPR008930; Terp_gyc_toroid.
DR InterPro; IPR008993; TIMP_like.
DR Pfam; PF00207; A2M; 1.
DR Pfam; PF01835; A2M N; 1.
DR Pfam; PF01821; ANATO; 1.
DR Pfam; PF01759; NTR; 1.
DR PRINTS; PR00004; ANAPHYLATOXN.
DR SMART; SM00104; ANATO; 1.
DR SMART; SM00843; C34SC1.
DR PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR PROSITE; PS01178; ANAPHYLATOXIN 2; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_1.
DR PROSITE; PS00189; NTR; 1.
KW Signal.
FT SIGNAL 1 22 Potential.
FT CHAIN 1264 1642 beta chain of cobra venom factor.
FT CHAIN 23 648 alpha chain of cobra venom factor.
SQ SEQUENCE 1642 AA; 184516 MW; 2A71B2BD6D12A3 CRC64;

Query Match 49.1%; Score 4224; DB 2; Length 1642;
Best Local Similarity 49.6%; Pred. No. 8.2e-243;
Matches 824; Conservative 324; Mismatches 478; Indels 36; Gaps 20;

Qy 9 LLLLLLTHPLALGSPMSIITPNTLRLESSETWVLEAHDAQGDVPVTVVHDPGKKLV 68
Db 8 LVAAALIGPGSHGALYTLIIPAVERTDEQLIVEAHGSDTPKQLDIFVHDPKQKT 67

Qy 69 LSSEKTVLTPATNHMGNTFTIPANRPFKSEKGRNKFVTVQATGTQV-VEKVVVLSQS 127
Db 68 LFQTRVDMNPAGMLVTPTEIEFA-KEVSTDSRQNYVWQVT-GPQVRLEKVVLSYQS 125

Qy 128 GYLFIQTDKTIYTPGSTVLYRIFTVNHKLVPGRVVMVNIENPEGIPVKQDLSLSONCLG 187
Db 126 SFLFIQTDKGIYTPGSTVLYRVFSDHNTSNKNTVIVEFQTPGILVSSNVD----- 179

Qy 188 VLPLSW--DIPELVNGQWKIPAYYENSPOQVFSTEFVEKVEYVLPSEFIVEPEKPYI 245
Db 180 -LNFFWPYNLPDLVSLGTWRIVAKYHSPEN-YTAYFDVRKYVLPSEFVRLQSEKPYI 237

Qy 246 YNEKGLVITITARFLYKVKVEGTAFFVFIQDGEORISLPESLKRIPIEDGSGWVLSRK 305
Db 238 DGNENFHVITARYLYGEEVGVAVLFGVKLDDAKSIPDLSLRIPIIDGDKATLKR 297

Qy 306 VLLDGVQNPRAEDLVKSLYVSATVILHSGSDMQAERSGIPVTSPVQIHFTKPYFK 365
Db 298 TFRS--RFPNLNELVHTLYASVTWVTESGDMVWVTEQSGIHIVASPYQIHFTKPYFK 355

Qy 366 PGMPPFDLMVFTVNPDPGSPAYRVPVAVQGEDTVQSITQGDGVAKLSINTHPQKLSITVR 425
Db 356 PGVPYELTVYVNPDPGSPAHPVYVSEAFHSM-GTLLSDGTAKILAIPLNAQSLPIVR 414

Qy 426 TKQQLSEAEQATRWQALPYSTVGNSSNNYLHLSVLRTELPGETLNVNLLRMDRAHEA 485
Db 415 TNHGDLFRERQATKSMTAYQTGGSGNYLFAVITSTEIKPGDNLPNFNVKGNANSLK 474

Qy 486 KIRYTYLIMNKGRLKAGRVREPGQDLVWLPISITDTDFPSFRLVAYVTLIGASQRE 545
Db 475 QIKYFYLLINKGKIFKVGQRPDRDQNLVMTNLHITPDLIPSFVFAYIQV----GNNE 530

Qy 546 VVADSWVDVKDSCVGLSVKSGQSDRQPVPGQOMTLKIEGDHARVWLVAVDKGVFL 605
Db 546 VVADSWVDVKDSCVGLSVKSGQSDRQPVPGQOMTLKIEGDHARVWLVAVDKGVFL 605

```

Db 531 IVADSVWVDVKTCMGTLVVK---GDNLIQMPGAAMKIKLEGDPGARVGLVAVDKAVVVL 587
Qy 606 NKNKLTOSKIVDVVEKADIGCTPGSGKDYAGVSDAGLTFTSSGQOQTAQAELOCPQ 665
Db 588 NDKYISQAKWITIEKSPGCTAGSGQNNLVGFEDAGLAUTTSNLTKRSAAKCPQ 647
Qy 666 AARRR-SVOLTEKRMKDVGVY-PKELRCCEDGNENPMRFSQORTRFISLGEAKKV 723
Db 648 ANRRRSVLLLSNASKAAEFQDQDLRCCEDVHENPMGVTCEKRAKYOQEDACKAA 707
Qy 724 FLDCNVITLRRQARASHGLARSNLEDIIAENIVSRSEPPESLWNVDEL-KEPP 782
Db 708 FLECRVYKVRDNGRESELFLARDNEDGFIANSIISRSDFPKSLWLTOKLTBEPN 767
Qy 783 KNGISTKLMMIFKDSITTWELLAVMSDKKICVADPFVTVQDFFIDRLRLYSVVRN 842
Db 768 SQGISSTKMSFYLRDGIITWVLAWSFTTKGICVAEPYIRVMKVFIDLOMPYSVVK 827
Qy 843 EQVEIRAVLYNRYQONELKVRVELLNPAFCSLATTIKRRHOOTITIPPKSLVSPYVIVP 902
Db 828 EQVEIRALHNY-VNEDIYRVELLNPAFCASATKQRYQOPFIKALSSRAVFFVIVP 886
Qy 903 LKGLQVEVKAAYVHHFISDGRKSLKVVPEGRINKTVAVRULDPRLGRBQVQKEDI 962
Db 887 LEQGLHDEIKASQVQALMSDGRKSLKVVPEGVQ-KSIVTIVKLDPRKGVGTQLEVI 945
Qy 963 PPADLSQVDETSERITLQGPVQAQMTEDAVDAERLKHIVTPSGGCGQNMIGMPTV 1022
Db 946 KARKLDRVDETEIKIILQDPAQIENSIDGSKLNHLIITPSGGCGQNMIRMAVP 1005
Qy 1023 IAHVYLDSTQWKEFLKRGQMALEIKGYTQQLAFRQPSAFAAFVRAVPSWLTAYV 1082
Db 1006 IATYLDSTQWKEFLKRGQMALEIKGYTQQLAFRQPSAFAAFVRAVPSWLTAYV 1065
Qy 1083 VKVPSLAUNLIA-IDSQVCGAVKVLILEKQPDGVQFQEDAPVHQEMIGLRNNKNNK 1141
Db 1066 VKVPSLAUNLIA-IDSQVCGAVKVLILEKQPDGVQFQEDAPVHQEMIGLRNNKNNK 1124
Qy 1142 ALTAFLVLSIQEAKDICEQVNSLPQSIKAGDFLEANYMMLORSYTVIAGYALAOQMR 1201
Db 1125 YLTAFLVLSIQEAKDICEQVNSLPQSIKAGDFLEANYMMLORSYTVIAGYALAOQMR 1184
Qy 1202 LKGPLNKFTTAKDKNRWEDPGKQLYNVEATSYALLALLOLQKDFVFPVVRVNLQRY 1261
Db 1185 LNDRLVMAASTGRD-HWEYNAHTNIEGTSYALLALLOLQKDFVFPVVRVNLQRY 1242
Qy 1262 YGGYGSQTATFVFOALAQYKQADPHQELNLDVSLQIPSRSKITHRIHWSASLLRS 1321
Db 1243 YGETYGTQATVAFQALAEYELQMPHKLNDLITIELPREVPIRYRINYENALLART 1302
Qy 1322 EETKENGFTVTAEGKQGTLSVVTMYHAKAQDL-TCKNFDLKVITKPAPEETKRPQDA 1380
Db 1303 VETKLNODITVTAEGKQGTLSVVTMYHAKAQDL-TCKNFDLKVITKPAPEETKRPQDA 1359
Qy 1381 KNTWMLICRYRGDODATMSILDISMTGEPADDDLLKOLANGVDRVYSKYVELDKAFSD 1440
Db 1360 KGALMLKICRYLGEVDSTWIIIDISMLTGFPLDIEDLITLSKGVDRVYSKYVELDKAFSD 1419
Qy 1441 RNTLIYKLVSHSDDCLAFKVQVFNVELIQGAVKVIAYYNLEBSCFTFYHPEKEDG 1500
Db 1420 KVAVILYLVSHSDDCLAFKVQVFNVELIQGAVKVIAYYNLEBSCFTFYHPEKEDG 1479
Qy 1501 KLNKLCRDELCAEENCFQKSDDKVTEERLDCACEPGVDYVYKTRLVKVLNSDFDE 1560
Db 1480 LLNKICIGNVCRGAGETCSLNRHQBIDVPLQIEKACETNVYVYKTRLVKVLNSDFDE 1539
Qy 1561 YMAIEGTIISGSDEQVQOQRTFISPIKCRALXLEBKHYLMWGLSSDFWQKPNLSY 1620
Db 1540 YVMDVLEVIQGTDENFRKTHQYISQKQCBALNKNVDYLLWGSRLLPKDKISY 1599
Qy 1621 IIGKDTVWEHPDEDECOBENKQOQDLGAFETSNWVFGCP 1662
Db 1600 IITKNTWIERHPDEDECOBENKQOQDLGAFETSNWVFGCP 1641

RESULT 11
Q91588 PRELIMINARY; PRT; 1589 AA.
ID Q91588
AC Q91588;
DT 01-NOV-1996 (TREMUREL.01, Created)
DT 01-NOV-1996 (TREMUREL.01, Last sequence update)
DT 01-MAR-2004 (TREMUREL.26, Last annotation update)
DE Complement component C3 (Fragment).
OS Xenopus laevis/gilli.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_taxid=8359;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=95180322; PubMed=7875221;
RA Lambiris J.D., Pappas J., Mavroidis M., Wang Y., Manzone H.,
Schwager J., Du Pasquier L., Silibovsky R.,
RT "The third component of xenopus complement: cDNA cloning, structural
and functional analysis, and evidence for an alternate C3
transcript.";
RL Eur. J. Immunol. 25:572-578(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Lambiris J.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U19253; AAB60608.1; -;
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004866; P:endopeptidase inhibitor activity; IEA.
DR GO; GO:0006956; P:complement activation; IEA.
DR GO; GO:0006954; P:inflammatory response; IEA.
DR InterPro; IPR002890; A2M.N.
DR InterPro; IPR009048; AM_receptor_bind.
DR InterPro; IPR000020; Anaphylatoxin.
DR InterPro; IPR001840; Anaphylatoxin.
DR InterPro; IPR008964; Invasin_inclimin.
DR InterPro; IPR001599; Macroglobulin2.
DR InterPro; IPR001134; Netrin_C.
DR InterPro; IPR008930; Terp_cyc_toroid.
DR InterPro; IPR008993; TIMP_like.
DR Pfam; PF00207; A2M; 1.
DR Pfam; PF01835; A2M.N; 1.
DR Pfam; PF01821; ANATO; 1.
DR Pfam; PF01759; NTR; 1.
DR PRINTS; PR00004; ANAPHYLATOXN.
DR ProDom; PD003264; Anaphylatoxin; 1.
DR SMART; SM00104; ANATO; 1.
DR SMART; SM00643; C345C; 1.
DR PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
DR PROSITE; PS50189; NTR; 1.
FT NON_TER 1.
FT CHAIN 592 1489 complement component C3 alpha chain.
FT CHAIN <1 588 complement component C3 beta chain.
SQ SEQUENCE 1589 AA; 177904 MW; DCB777FB4B11456A CRC64;
Query Match 48.9%; Score 4206.5; DB 2; Length 1589;
Best Local Similarity 51.3%; Pred. No. 8.6e-242;
Matches 817; Conservative 288; Mismatches 465; Indels 23; Gaps 18;
Qy 83 MGNVTFIPANRFEKSGKRNKVTTCATGTCQVTEKVVLSQSGYFIQDTKITYPG 142
Db 8 LGTAVKIPSGDSKDTKSK-QVYVYVSKNVCALEKVVLSYHSGYIFLQDTKITYPG 66
Qy 143 STVLRYFTVNHKLLPVGRVTVNINENPEGIVKQDSLSQNLGVFLSWDIPELNMG 202
Db 67 STVLRYIVSMYKQPIKSLTIEFTPDGIIIVNTHLQDSKGIISQSHKLPPELLSLG 126

AC QSVIAT;
 DT 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
 DB Complement C3-S.
 OS Cyprinus carpio (Common carp).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Cyprinus.
 OX NCBI_TaxID=7962;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hepatopancreas;
 RX MEDLINE=20203443; PubMed=10741402;
 RA Nakao M., Mutsuro J., Obo R., Fujiki K., Nonaka M., Yano T.;
 RT "Molecular cloning and protein analysis of divergent forms of the
 RT complement component C3 from a bony fish, the common carp (Cyprinus
 RT carpio): presence of variants lacking the catalytic histidine";
 RL Eur. J. Immunol. 30:858-866(2000).
 DR EXBL; AB016213; BAA36621.1; -;
 DR HSPB; P01026; I00F.
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0004866; F:endorpeptidase inhibitor activity; IEA.
 DR InterPro; IPR002890; A2M N.
 DR InterPro; IPR009048; AM Receptor bind.
 DR InterPro; IPR000020; Anaphylatoxin.
 DR InterPro; IPR008964; Invasin inhibitor.
 DR InterPro; IPR001599; MacroglobulinA2.
 DR InterPro; IPR001134; Netrin_C.
 DR InterPro; IPR008930; Tetrp_cyc_coroid.
 DR InterPro; IPR008993; TIMP_like.
 DR Pfam; PF00207; A2M; 1.
 DR Pfam; PF01835; A2M N; 1.
 DR Pfam; PF01821; ANATO; 1.
 DR Pfam; PF01759; NTR; 1.
 DR ProDom; PD003264; Anaphylatoxin; 1.
 DR SMART; SM00104; ANATO; 1.
 DR SMART; SM00643; C34SC; 1.
 DR PROSITE; PS00477; ALPHA 2 MACROGLOBULIN; 1.
 DR PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
 DR PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
 DR PROSITE; PS00189; NTR; 1.
 SQ SEQUENCE 1646 AA; 184194 MW; BB40BE1DE08A9390 CRC64;
 Query Match. 40.9%; Score 3521; DB 2; Length 1646;
 Best Local Similarity 44.1%; Pred. No. 8e-201;
 Matches 734; Conservative 323; Mismatches 569; Indels 38; Gaps 24;
 9 LLLLLLTHLPALGSPWYIIIPNLRLESEETVLEAHDAGQ-DVPVTVTVHDFPGKLL 67
 8 LTVVLSSPLLTLCDPLFVLSAPNLRVGSSENVFEAQDYGAAIEVKIIVKHPKDR 67
 68 VLSEKTVLTPANMGVNTFIPANREPKSEKGRNK-FVTVOATFGTVQVKEVVLVSLQ 126
 68 EILSQSVRLTADNNFQIIRKIPDDQNYFSDNPLEKQYVLOAQFPSTLEKEVMVSFQ 127
 127 SGVLFQTDKTIITPGSTVLYRIFTVNHKLLPWGRV-VAMNIENPEGIPVQDSLSQ 185
 128 SGVIFVOTDRPIITPASTVQYRIFSLTPNLEPLSQSGITVEIMNPQGITVSEKIFPVK- 186
 186 LGVLPISWDIPELVNNQWIKIRAYENSQQVFSTEFVEKVELVLPSEFVIVEPEKYYI 245
 187 -GMKSGKYPFEIASGIMKVVTLFNSNTPEKKFTADFEVKEVLPSTFEVKLKPSSFFV 245
 246 YNEKGLVETITARFLYKCKVEGTAFVIFGIQDGEORISLPSLKRPIEDSGSGWLSRK 305
 246 -GDPSTLVDEIAKYLFGQVDGNAFVVGVMNGEKKSIFPTSLQVQVIRREGTAELTSG 304
 306 VLLDGVGNPRAEDLVGKSLVNSNTVILHSGSDMVAERSGPIVTSFYIHTTKYKFX 365
 305 MITKTF--PNINLVGQSIYVSVLSTSGSEKVAERRGQIVTSPYSIHFKKTIOTFFK 362
 366 PGMPFDLMVFTNFDGSPAYRVPVAVQG-EDTVQSLQDGDGVAKLSINTHPSQKPLSITV 424

DB 363 PGMPFDVSVVYVTPDQTPAVNVEVVGSGQTVKGOTKINGIAKVTVNTLGSSSTOEITA 422
 QY 425 RTKQQLSEABQATRTMQALPY-STVGNNSNYHLVSLVLELTELPGSETLVNVLMDRAH 483
 DB 423 KTKQPOLRDNQAVKQVTAHAYTPKASKKNYJHIGIDAAELQMGDSMKV--FLMTGOSP 480
 QY 484 EAKRYTYTILMKNKRLKAGRQVRPBGDLVVLPLSITTTDFIPSPRLVAYTYLIGASQ 543
 DB 481 GVXQDITYMILSKGQIVSVDRFKER-GQSLVSLVPVTKVMVPSFVAYYHV---GS 535
 QY 544 REVVADSVWDVVDKDSVUGSLVWKSQSEDRQPY--PQQWTLKIEGDHARVVLVAVDKG 601
 DB 536 SEVVSDSVWVDKDTCKMLQVQ---VKDKQNTYDTGDEVKLEITGDPKAKVGLVWVDKA 592
 QY 602 VFVLMKNKLTQSKIWDVWEKADIGCTPGSGKDYGAVFSDAGLFTFTSSSGOOTAOAEIQ 661
 DB 593 VQVLN-KNRLTQTQIWDVIEKHDTGCTAGGRDRMGVFTDAGLMFQSTAGTNTIRTMPE 651
 QY 662 CPQPAARRRSVQLTERKMDKVKYKPKELKCCEDGRENPMRPFSCQRTIRISIGEACK 721
 DB 652 CPKPKRRRAESLLKTTSTLAEKYGELKQCCVGDGRENKLGYTCDRRSRVIADGKECV 711
 QY 722 KVFDCCNQYITELRRQCHARASHGLARSNLDDEIIAE-ENIVSRSEFPSPSWLVNVEDLKE 780
 DB 712 DAFLHCCNQMKTKDVKDEVEVMVARSDDDDDYTESSEIYSTRTOFPESMLWEEIDLCD 771
 QY 781 PPKNGISTKLNNVIFLKDSTITWILAVSMDKXGICVADPFVTVWQDFIDRLRPSYV 840
 DB 772 --KCAIPTKEKAIYVKDSITWQILAISLPTTGICVABEIEIVFKSLFIDLMKPYSAV 829
 QY 841 RNEQVEIRAVLYNRQNELKVRVELLHNPAGFSLATTKRRHQQTITIPKSSLSVPYVI 900
 DB 830 RGEQLERAILHNVTPKKQ-KRVFEFMEDEDVCSASKKGKRYTIVSVDKSGSISVSVI 888
 QY 901 VPLKTGLQEVKKAAYTHHFIISDGVKSLKVPPEGIRMNKTAVAVRTLDERLGRGV-QK 959
 DB 889 IPMTLGNHMEVFKASAYDSIYTDGVRKPLKVAEGVLVPLQRONLELNPAKNGEPILIK 948
 QY 960 EDIPPADLSQDVPDTESETRILQGTFFVAQMTDAVDAERLKHLYVTPSGCGEQNMIGT 1019
 DB 949 GEIP---ADRPDTPANTVISITGEBISQVBOAISGDFMGKLIYQPSGCGEQNMIGT 1004
 QY 1020 PTVTAVHVLDETEQWEXFLEKQGALELIKKGQYTOCLAFROPSSAFARVBAESTWLT 1079
 DB 1005 LPVATHYLDNTQWNAVGIQRTTEALNHIRTGYQLTYRKSDGSGYAAFTSRPSTWLT 1064
 QY 1080 AYVVKVSLAVNLIAIDSQVLCGRVWLILEKQKPDGVPFOEDAPVTHQEMIGGLRNNEK 1139
 DB 1065 AYVAKVPAIANNVTIEENVICSAKWLVLHKQLPDGSPKEDSAVSSGWDGVDQGN-A 1123
 QY 1140 DMALTAFVLISLOAKDICEQVNSIPGSITKAGDFLEANYMKNLQSRVYVATAGVALAQ 1199
 DB 1124 DASLTAFVVIAMQGEIEICAGSVASQESIRKAVSLEGRLPQLTNPYAVAMTSVAMASE 1183
 QY 1200 GRKGLPLNKFLTTAKDKRWEDFGKOLYNVEATSALLALLQLKDFDVPVFWRLNMQ 1259
 DB 1184 NKLKNDILMRHSTOQESASRGMTVPVQHHLSLEATAVALVRDRDFDAGEAVHWLARQ 1243
 QY 1260 RYVGGVGSCTAFTWFCALAQYKADPHQELNVLVSLQLPSSRSKITHRIHWSASLL 1319
 DB 1244 QAHYGGYGTQTATLMVFQAVAEYNTQVKTQNFNLDELVEISVAGRSKPVWTFKRDNAHL 1303
 QY 1320 RSEETKENEGFTVTAEGKGQGTLSVWIMYHAK-AKQDLTCNKPDLKVTIKPAPETEKRQ 1378
 DB 1304 RSDRVEINKDFSVTARGTKATLSVLTYYARPAVKSDCTLLFDL--TVKMEEDNKSNOE 1361
 QY 1379 DAKVTMILEICTRYRGP-QDATMSILDISMTGTFAPDQDLKQLANGVDRYTSKYELDKA 1437
 DB 1362 GAIEYKFTMPFYKSKKKAATWILLDGIPTGFSVDSRDLQELSTGKERYIQKFEMDKV 1421
 QY 1438 FSDRNTLIYLDKVSHEDEDCLAPKHQYVFNVELIFGAVKVVAYVYNNLSESCIRFVHPEK 1497

Db 1422 LSEGLILYLEKVLHREKERIAFRMHKIQNVGLLOPAAVTIYEYSPDARCTKPPHPR 1481
 Qy 1498 EDGLKLNKLDCLCRCAENCFOKSDDKVTLBERLDKACEPGVDVYKTRLVKVLQSLND 1557
 Db 1482 EDGALYRUCKDMCCCAENCSYOKK-NHVEDDERFNKACEAGDVIYKVVGMGLKKD 1540
 Qy 1558 FDEYMAEQTKSGSDEVQOQRTTISPICKREALKLEKKHYLMWGLSLDFFWGEKEN 1617
 Db 1541 SDIEMKVEQVLSKGTDEBEGVTRTELARPSCREHLGLVKGSYLIMGKSVLPLKJGS 1600
 Qy 1618 LSYIIGKDTVWEHPEDECOENOKQCDLGAFTESVVFQC 1661
 Db 1601 LOYINGEQTWYPTREESQTRHRRDYIGISDLENSLRKEC 1644

RESULT 15
 Qy9IA8 PRELIMINARY; PRT; 1640 AA.
 AC Q9YIA8;
 DT 01-MAY-1999 (TremBLrel. 10, Created)
 DT 01-MAY-1999 (TremBLrel. 10, Last sequence update)
 DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
 DE Complement C3-H2.
 OS Cyprinus carpio (Common carp).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Cyprinus.
 OX NCBI_TaxID=7962;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hepatopancreas;
 RX MEDLINE=20203443; PubMed=10741402;
 RA Nakao M., Mutsuro J., Obo R., Fujiki K., Nonaka M., Yano T.;
 RT "Molecular cloning and protein analysis of divergent forms of the
 RT complement component C3 from a bony fish, the common carp (Cyprinus
 RT carpio): presence of variants lacking the catalytic histidine.";
 RL Eur. J. Immunol. 30:858-866(2000).
 DR EMBL; AB016212; BAA36620.1; -;
 DR HSSP; P01026; 1QQF.
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
 DR GO; GO:0006956; P:complement activation; IEA.
 DR GO; GO:0006954; P:inflammatory response; IEA.
 DR InterPro; IPR002890; A2M_N.
 DR InterPro; IPR003048; AM_Receptor_bind.
 DR InterPro; IPR000020; Anaphylatoxin.
 DR InterPro; IPR001840; Anaphylatoxin.
 DR InterPro; IPR001599; Macroglobulin2.
 DR InterPro; IPR001134; Netrin_C.
 DR InterPro; IPR008930; Terp_cyc_toroid.
 DR InterPro; IPR008993; TIMP_like.
 DR Pfam; PF00207; A2M; 1.
 DR Pfam; PF01835; A2M_N; 1.
 DR Pfam; PF01821; ANATO; 1.
 DR Pfam; PF01759; NTR; 1.
 DR PRINTS; PR00004; ANAPHYLATOXN.
 DR ProDom; PD003264; Anaphylatoxin; 1.
 DR SMART; SM00104; ANATO; 1.
 DR SMART; SM00643; C345C; 1.
 DR PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
 DR PROSITE; PS01177; ANAPHYLATOXIN_1; UNKNOWN_1.
 DR PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
 DR PROSITE; PS0189; NTR; 1.
 SQ SEQUENCE 1640 AA; 183695 MW; 49EB0845275BF159 CRC64;

Query Match 40.5%; Score 3489; DB 2; Length 1640;
 Best Local Similarity 43.7%; Pred. No. 6.5e-199;
 Matches 726; Conservative 327; Mismatches 568; Indels 42; Gaps 22;

Qy 9 LLLLLLTHLPALGSPHYSITPNIKLSEETNVLEAHD-AQGVPVTVTHDPGKGL 67
 Db 8 LTVLLSPLLTCDPLVLSAPNLLRVGSSNNFVRAQYSGEDLNKXIIIVNHPKXDM 67

Qy 68 VLSEKTVLTPATNHMGNVTFTIPANREPKSEKGRNK-FVTVOATFGTQVVEKVVLSLQ 126
 Db 68 DILSKSVLTAANCFOILDTIDKIPDDQNFSDPLEKQVYVLOAOPSPVTEKVVLSLQ 127
 Qy 127 SGYLFQTDKTIYTPGSTVLYRIFTVNHKLLPVGR-TVMVNIENPEGIPVKODSLSSNQ 185
 Db 128 SGYIFQTDKTIYTPGSTVLYRIFTVNHKLLPVGR-TVMVNIENPEGIPVKODSLSSNQ 186
 Qy 186 LGVLPISWDIPELVNMGOKIRAYIENSPOQVSTFEFEVKEVLPSPFVIEPTKFFYI 245
 Db 187 -GKSGNYALPEVASGIIWKLVTLFSNTEQKFTADFEVKEVLPFTFEVKLPSPSPFYV 245
 Qy 246 YNKGLEVTITAPFLYKVKVEGTAFVIFGIQDGEORISLPSSLRIPEDSGEVEVLGRK 305
 Db 246 -REQSLTVDIKAYLFGQNVDTAFVYFGWMENKXTRIPASLOKXKIMKEGTAELTNQ 304
 Qy 306 VLLDGVNPRAEADLVGKSLVVSATVILHSGSDMVAERSGIPVTSPIQIIFTTKPKYFK 365
 Db 305 MITKTP--PNINQLVGRSIVSVLSLTSSESEWAEARRGIOIVTSPYTHFRKTPHPFK 362
 Qy 366 PGMPFDLMVFTNPDGSPAYRVPVAVOGEDTVQSITQDGVAKLSINTHPQKPLSIVR 425
 Db 363 PGMPFVSIVYTNPDQTPAVRMEVEV-NPGGVKGRTRDNGIAKVMVNTPGSSTLEITAK 421
 Qy 426 TKQELSEAEQARTMQALPYSTVGNSSNYHLHSLVRLTELRDGETLVNVLRLMDRAHEA 485
 Db 422 TKDPKLDDQQAQVETWTAQAYIPKOGSSNYLHIGIDAAELOIGDSMTVN--LNTQIIPGV 479
 Qy 486 KIRYYTYLIMKGRLLKAGRQVREPGQDLVPLSITDFTPSFLVAYYILIGASQRE 545
 Db 480 TQDFTYMLSKGQIVKAVRFRKQ-QQSLVTLVSFVTKDMSPFRFAYYHV---GSSE 534
 Qy 546 VVADSVVMDVKDCVGLSVVVGSGQSDRQFV-PGQWMLTKIEDGHCAVVLVAVDKGVFV 604
 Db 535 VVSDSIWVDKDCMCTGLQIEVKHEKINTYGTGNKVELQITGDFGAKVGLVVDKAVQV 594
 Qy 605 LNKKNKLTQSKIWDVVEKADIQCTPGSGKDYAGVTSFSDAGLTTSSSSGQQTQARABLCQPQ 664
 Db 595 LN-KNRLTQTKIWDVIEKHDTGCTAGGSDSGVFTDAGLMEVSTAGCTNTRTPPECFI 653
 Qy 665 PAARRRSVQLTEKRMKVKYKPKELKCKCEQDMENPMRPFQCORRTFISLGEACKVVF 724
 Db 654 HAKKRAENLLQITGLAGKSGELKQCCVQDGMKLGITCERRAMVLDGECVRAF 713
 Qy 725 LDCNVITELRROHARASHLGLARSLNLD--IAENIVSRSEFPESFWNVNVDLKEP 781
 Db 714 LDCNCQMKRKTETEEBEMILARSDDDDDDDYDTSEBTSRTKFPESWLEVEYDLPTS 773
 Qy 782 PKNGISTKLMNIFLKDSTITWELAVSMGDKKIGICVADPFVTVMQDFFIDRLPYSVVR 841
 Db 774 DKGETTSIPKVIYKDSITTWQILAVLSLPTLIGICVAFPEEMAVFKQLFIDLKMPYSVR 833
 Qy 842 NEQVEIRAVLYNTRQNELKVRVELLNHPAFGLATTYRRHQQTITIPKSSLSVYVIV 901
 Db 834 GEQLEIKAIHNHYTPNKQ-KVRVELMKTEDVCSASKGKYRTVTVSDSGSSISYSYVII 892
 Qy 902 PLKTGLQVEVEKAAVYHHFISDGVRSKLVKVPGEIRMNKTVAVRTLDPERLREGVQKED 961
 Db 893 PMLTGNHDIKVASADVIYTDGVKALKVSEGV-----LTSVHRENVELNPKGKPF 945
 Qy 962 IPADLSD-QVDETESETEILLQCTPVAQMTEDAVDAERLKHILVTPSGCGEQNMIGWTP 1020
 Db 946 IPKADIQDNRPDTPANTYISITGEEITQTVEQAISGSGFMGRLIVQPSGCGEQNMIGWL 1005
 Qy 1021 TVIAVHVLDETEQWEKFGLEKEQCALELTKKGYTQQLAFROPSSAAFAVFRAPSTWLTA 1080
 Db 1006 PVIATHYLDSTQWETVGFERNEINHLRTGYQLGVKRSKSGSYAAWIHRPSTWLTA 1065
 Qy 1081 YVVKVSLVNLNLAIDSOVLCCAVKWLILEKQKPGVFCQEDAPVJHQEMIIGLRNNNEKO 1140
 Db 1066 YVAKVFAMANNLVITBEKVLCSALEWLVLRKTQDGSFKESAVIHGEMGVQV-GNDAD 1124
 Qy 1141 MALTAFLVLSLQBAKDCIEQVNSLPGSITKAGDFLEANYMNLQRSYTVTAIAGVALAQMG 1200

Db 1125 ASLTAFAVIAOFARICAKE--SLHERIRKAVAYLEGRLSQLTNPYAVALTSYAMANEN 1182
Qy 1201 RLKGPLLNKFLTTAKDNWEDPGKOLYNVEATSVALALLQLKDFDPVPVVRWLNQOR 1260
Db 1183 KLNKDILMKHSQRGAGRSWTVPQGHHSLEATAYAVLALVAKADFDKAGEAVHLNRQ 1242
Qy 1261 YGCGYSTQATMVPQALAOYQKADAPDQELNDVSLQSPRSKITHRIHWESALLR 1320
Db 1243 SHYGGYGTQSTIMVFAVAYRTQMKQOQNFNLEVELSVEGRSRPVRYTIKKANALTR 1302
Qy 1321 SEETKENEGFTVTAEGKGOTLSVVTMYHAKADOLT-CNKFPDLKVTIKPAPETEKRPD 1379
Db 1303 SDRVDINQENVTARGTATLSVLTLYARPVEKSDCTFNLTV-----KNEKYNKG 1356
Qy 1380 AKNTMILEICTRYRQDQ-DATMSILDISMIGFAPDTPDQLQLANGVDRIYSKYELDKAF 1438
Db 1357 AVASYKLTWDFIYKSDKTATNTILDIGLPTGFEVESDLKELSSGKERYIQKYENNKVL 1416
Qy 1439 SDSNTLIYYLDKVSHSEDDCLAFKVHOFNVVELIOPGAVKVYAYYNLEESCTREYHPEKE 1498
Db 1417 SERGSLILYLDKVSHKEKEVIFSRMHQMLDVGLLOPAAVTIYEYSPDARCRKPYHPERT 1476
Qy 1499 DGKLNKCRDELICRAENCFTQKSDDKVTLERLDKACEPGVDYVYKTRLVKVQLSNDP 1558
Db 1477 DGAIYRLCKGDLCYCAEENCYQKK-NRVSDEERENRACEAGMDYVYKTVVGMDLKQDS 1535
Qy 1559 DEVIMAEQOTIKSGSDEVOVGQORTFISPIKCREALKLEEKHLYMGLSSDPWGEKPNL 1618
Db 1536 DIYDMKVEQVLKEGTDVVEEKVRPFLARPSCRBYLGLLEKSYLNGRSVDLPRLRSL 1595
Qy 1619 SYIIGKDTWVHWPDEDEQDEENKQCCQDLGAFTESMVVFGC 1661
Db 1596 QYVFGQTWVYWPTEESSTREHRYNGIAKLNLSLLSYGC 1638

Search completed: December 22, 2004, 00:37:16
Job time : 253 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2004, 00:16:53 ; Search time 197 Seconds

(without alignments)
3028.258 Million cell updates/sec

Title: US-09-875-519A-22

Perfect score: 8609

Sequence: 1 MGTSFGSLLLLLLTHLPLA.....KQCQLGAFTEMVVFGCPN 1663

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Genesep_23sep04.*

1: Genesep1980s.*
2: Genesep1990s.*
3: Genesep2000s.*
4: Genesep2001s.*
5: Genesep2002s.*
6: Genesep2003as.*
7: Genesep2003bs.*
8: Genesep2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	8609	100.0	1663	2 AAW34606	AAW34606 Wild type
2	8605	100.0	1663	2 AAW34607	AAW34607 Human C3
3	8605	100.0	1663	2 AAW34610	AAW34610 Human C3
4	8605	100.0	1663	2 AAW34609	AAW34609 Human C3
5	8605	100.0	1663	2 AAW34616	AAW34616 Human C3
6	8604	99.9	1663	2 AAW34618	AAW34618 Human C3
7	8603	99.9	1663	2 AAW40989	AAW40989 Human C3
8	8601	99.9	1663	2 AAW34608	AAW34608 Human C3
9	8600	99.9	1663	2 AAR34028	AAR34028 Human C3
10	8599	99.9	1663	2 AAW34627	AAW34627 Human C3
11	8598	99.9	1663	7 ADB90023	ADB90023 House com
12	8598	99.9	1663	7 ADD93518	ADD93518 Novel NOV
13	8598	99.9	1663	8 ADK12322	ADK12322 Human com
14	8598	99.9	1663	8 ADN04780	ADN04780 Antipsori
15	8597	99.9	1663	8 AAW40988	AAW40988 Human C3
16	8596	99.8	1663	2 AAW34621	AAW34621 Human C3
17	8596	99.8	1663	2 AAW34620	AAW34620 Human C3
18	8594	99.8	1663	2 AAR34029	AAR34029 Human mod
19	8593	99.8	1663	2 AAW34617	AAW34617 Human C3
20	8592	99.8	1667	2 AAW34631	AAW34631 Human C3
21	8588	99.8	1663	2 AAW34619	AAW34619 Human C3
22	8588	99.8	1663	2 AAW34628	AAW34628 Human C3
23	8586	99.7	1663	2 AAW34613	AAW34613 Human C3
24	8581	99.7	1663	2 AAR34030	AAR34030 Human mod
25	8581	99.7	1663	2 AAW34611	AAW34611 Human C3

26	8579	99.7	1663	2 AAW34614	AAW34614 Human C3
27	8578	99.6	1663	2 AAW34615	AAW34615 Human C3
28	8571	99.6	1663	2 AAW34612	AAW34612 Human C3
29	8567	99.5	1667	2 AAW34626	AAW34626 Human C3
30	8566	99.5	1663	2 AAW34630	AAW34630 Human C3
31	8554	99.4	1657	2 AAW34629	AAW34629 Human C3
32	8548	99.3	1663	2 AAW40990	AAW40990 Human C3
33	8545	99.3	1661	2 AAW34625	AAW34625 Human C3
34	8452	98.2	1635	2 AAW34624	AAW34624 Human C3
35	8193	95.2	1592	2 AAW34623	AAW34623 Human C3
36	7834.5	91.0	1540	4 ABG25976	ABG25976 Novel hum
37	6825	79.3	1663	8 ADI82104	ADI82104 Complemen
38	6752	78.4	1663	7 ADB90024	ADB90024 Mouse com
39	6752	78.4	1663	8 ADK12323	ADK12323 Mouse com
40	6272	72.9	1255	6 ABR63374	ABR63374 Human Alz
41	4386.5	51.0	1651	2 AAY23222	AAY23222 Pre-pro-c
42	4224	49.1	1642	2 AAY23729	AAY23729 Cobra ven
43	4224	49.1	1642	5 AAE21961	AAE21961 Pre-pro c
44	4224	49.1	1642	5 AAB71433	AAB71433 Cobra pro
45	4224	49.1	1650	5 AAB71469	AAB71469 Cobra pro

ALIGNMENTS

RESULT 1
AAW34606
ID AAW34606 standard; protein; 1663 AA.
XX
AC AAW34606;
XX
DT 09-APR-1998 (first entry)
XX
DE Wild type human C3 protein.
XX
KW Human; C3 protein; convertase; complement pathway protein; infection;
KW down-regulation resistant C3 convertase; xenograft rejection; therapy;
KW complement-mediated disease; autoimmune disease; leukaemia cell; tumour;
KW complement-mediated response; MHC-mismatched lymphocyte.
XX
OS Homo sapiens.
XX
PN WO9732981-A1.
XX
PD 12-SEP-1997.
XX
PF 04-MAR-1997; 97WO-GB000603.
XX
PR 07-MAR-1996; 96GB-00004865.
PR 07-JUN-1996; 96GB-00011896.
PR 08-JUL-1996; 96GB-00014293.
PR 19-NOV-1996; 96GB-00024028.
XX
PA (IMUT-) IMUTRAN LTD.
XX
PI Farries TC, Harrison RA;
XX
DR WPI; 1997-457534/42.
XX
N-PSDB; AAT93749.
PT Modified complement pathway protein that forms C3 convertase resistant to
PT down-regulation - used to exhaust the complement pathway by super-
PT activation, especially for preventing graft rejection, etc.
XX
PS Disclosure; Fig 1; 123pp; English.
XX
CC This sequence represents the wild type human C3 protein. This protein can
CC be mutated to produce a protein of the invention. The protein of the
CC invention is a modified native complement pathway protein (A) that forms
CC a down-regulation resistant C3 convertase. (A), their variants, fragments
CC and conjugates are used to deplete levels of complement pathway proteins
CC (by superactivation until one or more components are exhausted),
CC specifically to prevent rejection of foreign material (particularly a

CC xenograft) but also to prevent complement-mediated diseases resulting
 CC from (surgical) injury or antibody-antigen interaction in autoimmune
 CC disease, also to localise and/or amplify endogenous complement protein
 CC conversion and deposition at a specific site (e.g. a virus, infected cell
 CC or tumour, to increase sensitivity to complement-mediated responses; a
 CC particular application is eliminating any cancer cells left after
 CC surgical removal of a tumour). Also contemplated is ex vivo treatment,
 CC especially by passing blood through a matrix containing (A) (this may
 CC remove additional anaphylactic peptides and other inflammatory mediators)
 CC or killing of leukaemia cells or MHC-mismatched lymphocytes in extracted
 CC bone marrow. Since (A) is not inhibited by factor I, it can bind
 CC repeatedly to factor B (which is then inactivated), causing inactivation
 CC of the alternative pathway by consumption of factor B
 XX

SQ Sequence 1663 AA;
 Query Match 100.0%; Score 8609; DB 2; Length 1663;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGTSGPSLLLLTHLPLALGSPMYSITINILRSEETMVLBAHDAQDVPVTVTH 60
 DB 1 MGTSGPSLLLLTHLPLALGSPMYSITINILRSEETMVLBAHDAQDVPVTVTH 60
 QY 61 DFGKCLVLSSEKTLTPATNMGNVTFTIPANREFKSEKGRNFVTVQATFGTVVEKV 120
 DB 61 DFGKCLVLSSEKTLTPATNMGNVTFTIPANREFKSEKGRNFVTVQATFGTVVEKV 120
 QY 121 VLVSLSQGLFQTDKTIITPGSTVLYRIFVNHKLLPVGTVMNINENPEGIPVKQDSL 180
 DB 121 VLVSLSQGLFQTDKTIITPGSTVLYRIFVNHKLLPVGTVMNINENPEGIPVKQDSL 180
 QY 181 SSQNLQGLPLSWDIPELVNMGNKIRAYENSPOQVSTEFVEKVEYVLPSEFVIVPE 240
 DB 181 SSQNLQGLPLSWDIPELVNMGNKIRAYENSPOQVSTEFVEKVEYVLPSEFVIVPE 240
 QY 241 KFYIYNKGLVETITARELYKKVEGTAFVIFGIDQGEQRIISLPESLKRIPEDSGSEV 300
 DB 241 KFYIYNKGLVETITARELYKKVEGTAFVIFGIDQGEQRIISLPESLKRIPEDSGSEV 300
 QY 301 VLSRKVLDDGVNPRADLVGSLVYSATVILHSGSDMVQAESGIPVTSPIQIHFTKT 360
 DB 301 VLSRKVLDDGVNPRADLVGSLVYSATVILHSGSDMVQAESGIPVTSPIQIHFTKT 360
 QY 361 PKYFKGMPFDMVFTNPDGSPAYRVPVAVOGEDTVOSLTQGDGVAKLSINTHPSQKPL 420
 DB 361 PKYFKGMPFDMVFTNPDGSPAYRVPVAVOGEDTVOSLTQGDGVAKLSINTHPSQKPL 420
 QY 421 SITVTRTKQELSEAEQATMQLPYSITVGNSSNNYVLSVLTRELPGETLVNVLRLMD 480
 DB 421 SITVTRTKQELSEAEQATMQLPYSITVGNSSNNYVLSVLTRELPGETLVNVLRLMD 480
 QY 481 RAHEAKIRYTYLIMNKGSLKAGQVBPQGLVPLSTITTPISPRVAYYTLIGA 540
 DB 481 RAHEAKIRYTYLIMNKGSLKAGQVBPQGLVPLSTITTPISPRVAYYTLIGA 540
 QY 541 SGQREVVADSVVDVSDKSCVGLSVKSGQSEDRQVPVQCGQMTLKEGDHGARVVLVAVDK 600
 DB 541 SGQREVVADSVVDVSDKSCVGLSVKSGQSEDRQVPVQCGQMTLKEGDHGARVVLVAVDK 600
 QY 601 GVEVINKKILQSKITWDVVEKADIGCTPGSKDQVAGVPSDAGLTFTSSSGQQAQRAEL 660
 DB 601 GVEVINKKILQSKITWDVVEKADIGCTPGSKDQVAGVPSDAGLTFTSSSGQQAQRAEL 660
 QY 661 QCQPAARRRRSRYLTKEKMDKVGYPKELRKCCEGDMRENPMRFSQQRTRFISLGEAC 720
 DB 661 QCQPAARRRRSRYLTKEKMDKVGYPKELRKCCEGDMRENPMRFSQQRTRFISLGEAC 720
 QY 721 KVFLLDCCNYITELRQHARASHLGLARNLDEDEIIAENIVSRSEPPESLWLNVEDLKE 780
 DB 721 KVFLLDCCNYITELRQHARASHLGLARNLDEDEIIAENIVSRSEPPESLWLNVEDLKE 780
 QY 781 PPKNGISTKLMNIFLKDSITTWELAVSMDSKKGICVADPFVTVMQDFFIDLELPYSV 840

RESULT 2
AAW34607

ID AAW34607 standard; protein; 1663 AA.

XX AAW34607;

XX AC AAW34607;

XX DT 09-APR-1998 (first entry)

XX DE Human C3 protein mutant C3M-123.

XX KW Human; C3 protein; convertase; complement pathway protein; infection;
 KW down-regulation resistant C3 convertase; xenograft rejection; therapy;
 KW complement-mediated disease; autoimmune disease; leukaemia cell; tumour;

complement-mediated response; MHC-mismatched lymphocyte; mutein.

XX Homo sapiens.

OS Key Location/Qualifiers

PH Misc-difference 1303

FT /note= "Arg to Gln mutation"

XX WO9732981-A1.

XX 12-SEP-1997.

XX 04-MAR-1997; 97WO-GB0006503.

XX 07-MAR-1996; 96GB-00004865.

XX 07-JUN-1996; 96GB-00011896.

XX 08-JUL-1996; 96GB-00014293.

XX 19-NOV-1996; 96GB-00024028.

XX (IMUT-) IMUTRAN LTD.

XX Farries TC, Harrison RA;

XX WPI; 1997-457534/42.

XX Modified complement pathway protein that forms C3 convertase resistant to

PT down-regulation - used to exhaust the complement pathway by super-

PT activation, especially for preventing graft rejection, etc.

XX Example 4; Page; 123pp; English.

XX This sequence represents a mutated human C3 protein of the invention (see

CC AAW34606 for wild type protein). This protein is a protein of the

CC invention, and is a modified native complement pathway protein (A) that

CC forms a down-regulation resistant C3 convertase. (A), their variants,

CC fragments and conjugates are used to deplete levels of complement pathway

CC proteins (by superactivation until one or more components are exhausted),

CC specifically to prevent rejection of foreign material (particularly a

CC xenograft) but also to prevent complement-mediated diseases resulting

CC from (surgical) injury or antibody-antigen interaction in autoimmune

CC disease, also to localise and/or amplify endogenous complement protein

CC conversion and deposition at a specific site (e.g. a virus, infected cell

CC or tumour, to increase sensitivity to complement-mediated responses; a

CC particular application is eliminating any cancer cells left after

CC surgical removal of a tumour). Also contemplated is ex vivo treatment,

CC especially by passing blood through a matrix containing (A) (this may

CC remove additional anaphylactic peptides and other inflammatory mediators)

CC or killing of leukaemia cells or MHC-mismatched lymphocytes in extracted

CC bone marrow. Since (A) is not inhibited by factor I, it can bind

CC repeatedly to factor B (which is then inactivated), causing inactivation

CC of the alternative pathway by consumption of factor B

XX Sequence 1663 AA;

SQ

Query Match 100.0%; Score 8605; DB 2; Length 1663;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1662; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGPTSGPSLLLLTHLPLALGSPMYSIIIPNLRLSEETMVLBAHQDQGVPTVTVH 60

DB 1 MGPTSGPSLLLLTHLPLALGSPMYSIIIPNLRLSEETMVLBAHQDQGVPTVTVH 60

QY 61 DFPQKLVLSSEKTVLTPATNHMGNTFTIPANREFKSGRKNKFTVQATFTQVKEV 120

DB 61 DFPQKLVLSSEKTVLTPATNHMGNTFTIPANREFKSGRKNKFTVQATFTQVKEV 120

QY 121 VLVLQSGYLFIQTDKIYIPGSTVLRIPTVNHKLLPVGRVWNIENPEGIPVKQDSL 180

DB 121 VLVLQSGYLFIQTDKIYIPGSTVLRIPTVNHKLLPVGRVWNIENPEGIPVKQDSL 180

QY 181 SSQNLGLVPLSWDIPVLNMGQWKIRAYYENSFQVFTSTEFVKBYVLPSEVIVEPTE 240

DB 181 SSQNLGLVPLSWDIPVLNMGQWKIRAYYENSFQVFTSTEFVKBYVLPSEVIVEPTE 240

241 KFYIYNKGLVITITARELYGKKVEGTAFVIFGIDGQFORISLPESIKRPIEDSGSEV 300

241 KFYIYNKGLVITITARELYGKKVEGTAFVIFGIDGQFORISLPESIKRPIEDSGSEV 300

301 VLSRKVLDDGVQNPRAEDLVGKSLYVSATVILHSGSDMVQAERSGIPVTSYQIHFTKT 360

301 VLSRKVLDDGVQNPRAEDLVGKSLYVSATVILHSGSDMVQAERSGIPVTSYQIHFTKT 360

361 PKYFKGMPFDLMVFTVNDPDSPAYRVPVAVOGEDEVQSLTCGDGVAKLSINTHPSOKPL 420

361 PKYFKGMPFDLMVFTVNDPDSPAYRVPVAVOGEDEVQSLTCGDGVAKLSINTHPSOKPL 420

421 SITVRTKQELSEAEQATRTMQALPYSTVGNNSNNVHLVSLRTELRLPGETLVNAPFLRMD 480

421 SITVRTKQELSEAEQATRTMQALPYSTVGNNSNNVHLVSLRTELRLPGETLVNAPFLRMD 480

481 RAHEAKIRYTYVILNMKGRLLKAGQVREPQGLVVLPLSITTTDFIPFRVAVYTTIGA 540

481 RAHEAKIRYTYVILNMKGRLLKAGQVREPQGLVVLPLSITTTDFIPFRVAVYTTIGA 540

541 SGOREVVADSVWYVDKSDCVGSLVVKSGQSEDRQVPQGQMTLKIEGDHGARVVLVAVDK 600

541 SGOREVVADSVWYVDKSDCVGSLVVKSGQSEDRQVPQGQMTLKIEGDHGARVVLVAVDK 600

601 GVFVLNKKNLTKOSKIWDVVEKADIGCTPGSGDKYAGVFSADAGLTFTSSSGQOQAQRAEL 660

601 GVFVLNKKNLTKOSKIWDVVEKADIGCTPGSGDKYAGVFSADAGLTFTSSSGQOQAQRAEL 660

661 QCPQPAARRRRSVOLTEKMDKVGPKELRKCCDGMRENPMRPSCORRTFISLGEAC 720

661 QCPQPAARRRRSVOLTEKMDKVGPKELRKCCDGMRENPMRPSCORRTFISLGEAC 720

721 KKVFLDCCNYITELRQHARASHGLARSLNLDIEDIIAENIVSRSEFFESWLNVEDLKE 780

721 KKVFLDCCNYITELRQHARASHGLARSLNLDIEDIIAENIVSRSEFFESWLNVEDLKE 780

781 PPNNGISTKMLNIFLKDSTITWEILAVSMDSKKGICVADPPFEVTVMQDFFIDRLRPIYVV 840

781 PPNNGISTKMLNIFLKDSTITWEILAVSMDSKKGICVADPPFEVTVMQDFFIDRLRPIYVV 840

841 RNEQVIRAVLYNQRONQELKVRVLLHNPAFCSLATTKRHHQQTITTPKSSLSVPVVI 900

841 RNEQVIRAVLYNQRONQELKVRVLLHNPAFCSLATTKRHHQQTITTPKSSLSVPVVI 900

901 VPLKTLQEVKAAVYHHFISDGVKRSKLVVPEGIRKNTVAVRTLDPERLREGVOKE 960

901 VPLKTLQEVKAAVYHHFISDGVKRSKLVVPEGIRKNTVAVRTLDPERLREGVOKE 960

961 DIPPADLSQVDPDTESETRILLQGTVPVQOMTEDAVDAERLKHILVTPSGCGEONMIGMTP 1020

961 DIPPADLSQVDPDTESETRILLQGTVPVQOMTEDAVDAERLKHILVTPSGCGEONMIGMTP 1020

1021 TVTAVHYLDBETQWEKFGLEKQGALELIIKKGYTQCLAFRQPSFAFAFVKRAPSTWLTATA 1080

1021 TVTAVHYLDBETQWEKFGLEKQGALELIIKKGYTQCLAFRQPSFAFAFVKRAPSTWLTATA 1080

1081 YVVKVPSLANVLIATDSQVLGCAVKWVILEKQKPGVFEQEDAPVTHQEMIIGLRNNEKD 1140

1081 YVVKVPSLANVLIATDSQVLGCAVKWVILEKQKPGVFEQEDAPVTHQEMIIGLRNNEKD 1140

1141 MALTAFVLISLQSAKDICEEQVNSLPGSITKAGDFLEANYMNLQRSYTVIAGVALAQM 1200

1141 MALTAFVLISLQSAKDICEEQVNSLPGSITKAGDFLEANYMNLQRSYTVIAGVALAQM 1200

1201 RLKGPLLNKFLTTAKDKRNEWDPGKQLYVNEATSVALLALLQDKDFDVPVFWNLNEQR 1260

1201 RLKGPLLNKFLTTAKDKRNEWDPGKQLYVNEATSVALLALLQDKDFDVPVFWNLNEQR 1260

1261 YGGGKVGSTQATPMVFOALAQYOKADPHOEINLVDLSLQPSRSKITHRIHWESASLLR 1320

1261 YGGGKVGSTQATPMVFOALAQYOKADPHOEINLVDLSLQPSRSKITHRIHWESASLLR 1320

QY 1321 SEETKENEGFTVTAEGKGQGLSVVTMTHAKAKDQTCNKFDLKVTKIPAPETEKRPQDA 1380
 DB 1321 SEETKENEGFTVTAEGKGQGLSVVTMTHAKAKDQTCNKFDLKVTKIPAPETEKRPQDA 1380
 QY 1381 KNTMILEICTRYRGDQDATMSILDISMGTGFPADPTDDLKQLANGVDVRYISYELDKAFSD 1440
 DB 1381 KNTMILEICTRYRGDQDATMSILDISMGTGFPADPTDDLKQLANGVDVRYISYELDKAFSD 1440
 QY 1441 RNTLIIVLDKSHSEDDCLAFKQHVNFVELIQPGAVKYAYXNLEESCTRFYHPEKEDG 1500
 DB 1441 RNTLIIVLDKSHSEDDCLAFKQHVNFVELIQPGAVKYAYXNLEESCTRFYHPEKEDG 1500
 QY 1501 KINKLCRDELCAEENCFIQSDDKVLEERLDKACBPVDYVYKPLVKVQJNSDFDE 1560
 DB 1501 KINKLCRDELCAEENCFIQSDDKVLEERLDKACBPVDYVYKPLVKVQJNSDFDE 1560
 QY 1561 YIMAIETIKSGSDEYVQVQQRFTFISPIKREALKLEBKXHYLMWGLSSDFWGEKPNLSY 1620
 DB 1561 YIMAIETIKSGSDEYVQVQQRFTFISPIKREALKLEBKXHYLMWGLSSDFWGEKPNLSY 1620
 QY 1621 IIGKDTWVHWPDEDCOEENKQKQODLGAFTESMVVGCPN 1663
 DB 1621 IIGKDTWVHWPDEDCOEENKQKQODLGAFTESMVVGCPN 1663

RESULT 3

ID AAW34610 standard; protein; 1663 AA.

AC AAW34610;

DT 09-APR-1998 (first entry)

DE Human C3 protein mutant CV-2.

XX Human: C3 protein; convertase; complement pathway protein; infection;
 KW down-regulation resistant C3 convertase; xenograft rejection; therapy;
 KW complement-mediated disease; autoimmune disease; leukaemia cell; tumour;
 XX complement-mediated response; MHC-mismatched lymphocyte; muten.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 776 /note= "E776K mutation"

ET

XX W09732981-A1.

XX 12-SEP-1997.

XX 04-MAR-1997; 97WO-GB000603.

XX 07-MAR-1996; 96GB-00004865.

XX 07-JUN-1996; 96GB-00011896.

XX 08-JUL-1996; 96GB-00014293.

XX 19-NOV-1996; 96GB-00024028.

XX (IMUT-) IMUTRAN LTD.

XX Farries TC, Harrison RA;

XX WPI; 1997-457534/42.

XX Modified complement pathway protein that forms C3 convertase resistant to
 PT down-regulation - used to exhaust the complement pathway by super-
 PT activation, especially for preventing graft rejection, etc.

PS Example 14; Page; 123pp; English.

XX This sequence represents a mutated human C3 protein of the invention (see
 CC AAW34606 for wild type protein). This protein is a protein of the
 CC invention, and is a modified native complement pathway protein (A) that
 CC forms a down-regulation resistant C3 convertase. (A), their variants,

CC fragments and conjugates are used to deplete levels of complement pathway
 CC proteins (by superactivation until one or more components are exhausted),
 CC specifically to prevent rejection of foreign material (particularly a
 CC xenograft) but also to prevent complement-mediated diseases resulting
 CC from (surgical) injury or antibody-antigen interaction in autoimmune
 CC disease, also to localise and/or amplify endogenous complement protein
 CC conversion and deposition at a specific site (e.g. a virus, infected cell
 CC or tumour, to increase sensitivity to complement-mediated responses; a
 CC particular application is eliminating any cancer cells left after
 CC surgical removal of a tumour). Also contemplated is ex vivo treatment,
 CC especially by passing blood through a matrix containing (A) (this may
 CC remove additional anaphylactic peptides and other inflammatory mediators)
 CC or killing of leukaemia cells or MHC-mismatched lymphocytes in extracted
 CC bone marrow. Since (A) is not inhibited by factor I, it can bind
 CC repeatedly to factor B (which is then inactivated), causing inactivation
 CC of the alternative pathway by consumption of factor B
 XX
 SQ Sequence 1663 AA;

Query Match 100.0%; Score 8605; DB 2; Length 1663;

Best Local Similarity 99.9%; Pred No 0;

Matches 1662; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGPTSGPSLLLLTHLPALGSPMYSIITPNILRESEBETMWLEAHDAGQGPVTVTVH 60

DB 1 MGPTSGPSLLLLTHLPALGSPMYSIITPNILRESEBETMWLEAHDAGQGPVTVTVH 60

QY 61 DFPKKLVLSSEKTVLTPATNMGNTFTIPANREFKSEKGRNKFVTVQATFGTOVEKV 120

DB 61 DFPKKLVLSSEKTVLTPATNMGNTFTIPANREFKSEKGRNKFVTVQATFGTOVEKV 120

QY 121 VLVSQSGYLFQTDKTYITFGSTVLYRIPTVNHKLLPVGRVVMNENPEGIPVKQDSL 180

DB 121 VLVSQSGYLFQTDKTYITFGSTVLYRIPTVNHKLLPVGRVVMNENPEGIPVKQDSL 180

QY 181 SSKQNLGVLPLSWDIPBLVNMGMKIRAYVENSPPQVFSTPEPEKVEYLPSEVIVEPTE 240

DB 181 SSKQNLGVLPLSWDIPBLVNMGMKIRAYVENSPPQVFSTPEPEKVEYLPSEVIVEPTE 240

QY 241 KFYIYNEKGLVITITARFLYKKGVECTAPVIFGIGQSGQRISLPSLKRIPIEGSGEV 300

DB 241 KFYIYNEKGLVITITARFLYKKGVECTAPVIFGIGQSGQRISLPSLKRIPIEGSGEV 300

QY 301 VLSRKVLDDGVQNPRAEDLVKSLYSATVILHSGSDMVQAESGPIVTSPIQIHFTKT 360

DB 301 VLSRKVLDDGVQNPRAEDLVKSLYSATVILHSGSDMVQAESGPIVTSPIQIHFTKT 360

QY 361 PRYFKPGMPFDLMVFTNPDGSPAYRVPVAVQGEDTVQSLTQGDGVAKLSINTHPSQKPL 420

DB 361 PRYFKPGMPFDLMVFTNPDGSPAYRVPVAVQGEDTVQSLTQGDGVAKLSINTHPSQKPL 420

QY 421 SITVTKKQELSEAEQATRTMQLPYSTVGNNSNNYLHLSVLRTLPGETLNVNFLRMD 480

DB 421 SITVTKKQELSEAEQATRTMQLPYSTVGNNSNNYLHLSVLRTLPGETLNVNFLRMD 480

QY 481 RAHEAKIRYTYLIMNKRLLKAGQVRBQDLVPLPSITITDPIPSRLVAYTYLIGA 540

DB 481 RAHEAKIRYTYLIMNKRLLKAGQVRBQDLVPLPSITITDPIPSRLVAYTYLIGA 540

QY 541 SGQREVADSVWVDVSDSCVGLSVKSGQSEDRQPVFGQMTLKIEGDHGARVAVDX 600

DB 541 SGQREVADSVWVDVSDSCVGLSVKSGQSEDRQPVFGQMTLKIEGDHGARVAVDX 600

QY 601 GUFVLNKKKLTQSKIKDWVEKADIGCTPGSGDKYAGVFSACLITSSSGQTAQRAEL 660

DB 601 GUFVLNKKKLTQSKIKDWVEKADIGCTPGSGDKYAGVFSACLITSSSGQTAQRAEL 660

QY 661 QCQPPAARRRRSVQTEKRMDDKVGKPKELRKCCEDGMRENPMRPSQRRTRFISLGEAC 720

DB 661 QCQPPAARRRRSVQTEKRMDDKVGKPKELRKCCEDGMRENPMRPSQRRTRFISLGEAC 720

QY 721 KKVFLDCCNYITELARHARASHLGLARSLNLDDEITAEENIVSRSEFPSSMWNVDLKE 780

DB 721 KKVFLDCCNYITELARHARASHLGLARSLNLDDEITAEENIVSRSEFPSSMWNVDLKE 780

Db 721 KKVFLDCCNYITELRQHRARASHGLARSNLDEDI IAEENI VSRSEFPESWLNWVLDKE 780
 Qy 781 PPKGIGSKLWNIPLKOSITTEWILAVMSDKGICVADPEVTVQWDFIDRLRLPSYV 840
 Db 781 PPKGIGSKLWNIPLKOSITTEWILAVMSDKGICVADPEVTVQWDFIDRLRLPSYV 840
 Qy 841 RNEQVEIRAVLYNRQNELKVRVELLHNFAPCSLATTKRHHQQTITIPKSSLSVPYVI 900
 Db 841 RNEQVEIRAVLYNRQNELKVRVELLHNFAPCSLATTKRHHQQTITIPKSSLSVPYVI 900
 Qy 901 VPLKTGLQGEVKAAYVHHFTSDGVRKSLKVVPEGIRMNKTAVRVLDPERLREGVQKE 960
 Db 901 VPLKTGLQGEVKAAYVHHFTSDGVRKSLKVVPEGIRMNKTAVRVLDPERLREGVQKE 960
 Qy 961 DIPPADLSQVDPTESETRILLQGTTPVAQMTEDAVDAERLKLHIVTPSGCGEONMIGMP 1020
 Db 961 DIPPADLSQVDPTESETRILLQGTTPVAQMTEDAVDAERLKLHIVTPSGCGEONMIGMP 1020
 Qy 1021 TVIAVHYLDETEQWEKFGLEKROGALBELIKGYTOQLAFPOSSAFAAFVKRAPSPWLTA 1080
 Db 1021 TVIAVHYLDETEQWEKFGLEKROGALBELIKGYTOQLAFPOSSAFAAFVKRAPSPWLTA 1080
 Qy 1081 YVVKVFLSIAVNLIAIDSOVLGCAVKWLIILEKQKPDGVFORDAFVHQMIGLGRNNNEK 1140
 Db 1081 YVVKVFLSIAVNLIAIDSOVLGCAVKWLIILEKQKPDGVFORDAFVHQMIGLGRNNNEK 1140
 Qy 1141 MALTAFLVLSIQEAKDICEQVNSLPGSITKAGDFLEANTYMLQRYTVTAIAGYALAOAG 1200
 Db 1141 MALTAFLVLSIQEAKDICEQVNSLPGSITKAGDFLEANTYMLQRYTVTAIAGYALAOAG 1200
 Qy 1201 RLKGPLLNKFTTAKDKNRWEDPGQLYNNVEATSYALLALLQKDFDPPVVRWLNQOR 1260
 Db 1201 RLKGPLLNKFTTAKDKNRWEDPGQLYNNVEATSYALLALLQKDFDPPVVRWLNQOR 1260
 Qy 1261 YVGGYGSTQATEVVFQALQYQKADPHQELNLDVSLQPSSESKITHRIWESASLAR 1320
 Db 1261 YVGGYGSTQATEVVFQALQYQKADPHQELNLDVSLQPSSESKITHRIWESASLAR 1320
 Qy 1321 SEETKENEFTVTAEGKGGQGLSVVWYHAKAKDQLTCKNFKDLKVTIKPAPETERPQDA 1380
 Db 1321 SEETKENEFTVTAEGKGGQGLSVVWYHAKAKDQLTCKNFKDLKVTIKPAPETERPQDA 1380
 Qy 1381 KNTWILEICTRYRGDQATWSILDISMTGFPADTDLDKQANGVDVRYISKVELDKAFSD 1440
 Db 1381 KNTWILEICTRYRGDQATWSILDISMTGFPADTDLDKQANGVDVRYISKVELDKAFSD 1440
 Qy 1441 RNTLIILYDKVSHSEDDCLAPKHQYFNVELIQGAVKVYAYYNLEESCTRFYHPEKEDG 1500
 Db 1441 RNTLIILYDKVSHSEDDCLAPKHQYFNVELIQGAVKVYAYYNLEESCTRFYHPEKEDG 1500
 Qy 1501 KLNKL-CRDELRCRAEENCFLQSDDKVTLERLDKACEPGVDVYVYKVLQVLSNDFE 1560
 Db 1501 KLNKL-CRDELRCRAEENCFLQSDDKVTLERLDKACEPGVDVYVYKVLQVLSNDFE 1560
 Qy 1561 YMAIEQTIKSGSDEVQVQOQTTFISPKREALKLEKHYLMWGLSDFDNGEKNLSY 1620
 Db 1561 YMAIEQTIKSGSDEVQVQOQTTFISPKREALKLEKHYLMWGLSDFDNGEKNLSY 1620
 Qy 1621 IIGKDTWVHEHPEDECOENQKQCDLGAFTESVWVFGCPN 1663
 Db 1621 IIGKDTWVHEHPEDECOENQKQCDLGAFTESVWVFGCPN 1663

RESULT 4

AAW34609

ID AAW34609 standard; protein; 1663 AA.

XX AC AAW34609;

XX AC AAW34609;

DT 09-APR-1998 (first entry)

XX DE Human C3 protein mutant C3M-51.

XX DE

Human: C3 protein; convertase; complement pathway protein; infection;
 down-regulation resistant C3 convertase; xenograft rejection; therapy;
 complement-mediated disease; autoimmune disease; leukaemia cell; tumour;
 complement-mediated response; MHC-mismatched lymphocyte; mutein.
 Homo sapiens.
 Key Location/Qualifiers
 Misc-difference 1320 /note= "Arg to Gln mutation"
 WT09732981-A1.
 12-SEP-1997.
 04-MAR-1997; 97MO-GB000603.
 07-MAR-1996; 96GB-00004865.
 07-JUN-1996; 96GB-00011896.
 08-JUL-1996; 96GB-00014293.
 19-NOV-1996; 96GB-00024028.
 (IMUT-) IMUTRAN LTD.
 Parries TC, Harrison RA;
 WPI; 1997-457534/42.
 Modified complement pathway protein that forms C3 convertase resistant to
 down-regulation - used to exhaust the complement pathway by super-
 activation, especially for preventing graft rejection, etc.
 Claim 7; Page; 123pp; English.

This sequence represents a mutated human C3 protein of the invention (see
 AAW34609 for wild type protein). This protein is a protein of the
 invention, and is a modified native complement pathway protein (A) that
 forms a down-regulation resistant C3 convertase. (A), their variants,
 fragments and conjugates are used to deplete levels of complement pathway
 proteins (by superactivation until one or more components are exhausted),
 specifically to prevent rejection of foreign material (particularly a
 xenograft) but also to prevent complement-mediated diseases resulting
 from (surgical) injury or antibody-antigen interaction in autoimmune
 disease, also to localise and/or amplify endogenous complement protein
 conversion and deposition at a specific site (e.g. a virus, infected cell
 or tumour, to increase sensitivity to complement-mediated responses; a
 particular application is eliminating any cancer cells left after
 surgical removal of a tumour). Also contemplated is ex vivo treatment,
 especially by passing blood through a matrix containing (A) (this may
 remove additional anaphylactic peptides and other inflammatory mediators)
 or killing of leukaemia cells or MHC-mismatched lymphocytes in extracted
 bone marrow. Since (A) is not inhibited by factor I, it can bind
 repeatedly to factor B (which is then inactivated), causing inactivation
 of the alternative pathway by consumption of factor B

Query Match 100.0%; Score 8605; DB 2; Length 1663;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1662; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGTPSGPSLLLLTHLPALGSPMYSIITPNILRESEETMVLHAHQDQDVPVTVVH 60
 Db 1 MGTPSGPSLLLLTHLPALGSPMYSIITPNILRESEETMVLHAHQDQDVPVTVVH 60
 Qy 61 DFPCKLVLSSEKTVLTPATNHMGVNTFTIPANRFEKSEKGNKFTVQATFGTVVEKV 120
 Db 61 DFPCKLVLSSEKTVLTPATNHMGVNTFTIPANRFEKSEKGNKFTVQATFGTVVEKV 120
 Qy 121 VLVSQSGVLFQIDTKTITVPGSTVLYRIFTVNHKLLPVGRVWVNIENPEGIPVKQDSL 180
 Db 121 VLVSQSGVLFQIDTKTITVPGSTVLYRIFTVNHKLLPVGRVWVNIENPEGIPVKQDSL 180

Sequence 1663 AA;
 Query Match 100.0%; Score 8605; DB 2; Length 1663;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1662; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGTPSGPSLLLLTHLPALGSPMYSIITPNILRESEETMVLHAHQDQDVPVTVVH 60
 Db 1 MGTPSGPSLLLLTHLPALGSPMYSIITPNILRESEETMVLHAHQDQDVPVTVVH 60
 Qy 61 DFPCKLVLSSEKTVLTPATNHMGVNTFTIPANRFEKSEKGNKFTVQATFGTVVEKV 120
 Db 61 DFPCKLVLSSEKTVLTPATNHMGVNTFTIPANRFEKSEKGNKFTVQATFGTVVEKV 120
 Qy 121 VLVSQSGVLFQIDTKTITVPGSTVLYRIFTVNHKLLPVGRVWVNIENPEGIPVKQDSL 180
 Db 121 VLVSQSGVLFQIDTKTITVPGSTVLYRIFTVNHKLLPVGRVWVNIENPEGIPVKQDSL 180

QY 181 SSQNLGVLPLSDIPELAVNMGMKIRAYVYENSPOQVFTSEPEVKEYVLPSPEVIVPE 240
 Db 181 SSQNLGVLPLSDIPELAVNMGMKIRAYVYENSPOQVFTSEPEVKEYVLPSPEVIVPE 240
 QY 241 KFYIYINEKLEVTITARFLYGGKVBGTARVIGIQDGEORISLPESLKRPIEDGSGEV 300
 Db 241 KFYIYINEKLEVTITARFLYGGKVBGTARVIGIQDGEORISLPESLKRPIEDGSGEV 300
 QY 301 VLSRKVLIDGQVONPRAEDLVGKSLVYSATVILHSGSDMVQABRSIGIPIVTSFYQHFTKT 360
 Db 301 VLSRKVLIDGQVONPRAEDLVGKSLVYSATVILHSGSDMVQABRSIGIPIVTSFYQHFTKT 360
 QY 361 KYFKPGMPLMVFVNTNPGSPAYRVPVAVQGEDTVQSLTQDGDVAKLSINTHPSQKPL 420
 Db 361 KYFKPGMPLMVFVNTNPGSPAYRVPVAVQGEDTVQSLTQDGDVAKLSINTHPSQKPL 420
 QY 421 SITVTRTKQELSEAEQATRTMQUALPYSTVGNSSNNYLHLSVLRTLEPGETLNVNPLRMD 480
 Db 421 SITVTRTKQELSEAEQATRTMQUALPYSTVGNSSNNYLHLSVLRTLEPGETLNVNPLRMD 480
 QY 481 RAHAKIRYTYLLIMNKGRLKAGQVRSPGQDLVVLPLSITTDFTBFRVLVAYYTLGA 540
 Db 481 RAHAKIRYTYLLIMNKGRLKAGQVRSPGQDLVVLPLSITTDFTBFRVLVAYYTLGA 540
 QY 541 SGOREVWADSVWVDVKDSCVGSILVWKSQSGSDRQPVFGQOMTLKIEGSHGARVVLVADK 600
 Db 541 SGOREVWADSVWVDVKDSCVGSILVWKSQSGSDRQPVFGQOMTLKIEGSHGARVVLVADK 600
 QY 601 GFVLNKNKLTQSKINDVVEKADIGCTPGSGKDYAGVPSDAGLFTSSSGOQRAEL 660
 Db 601 GFVLNKNKLTQSKINDVVEKADIGCTPGSGKDYAGVPSDAGLFTSSSGOQRAEL 660
 QY 661 QCPQPAARRRSVOLTSEKMDKVGYPKSLRKCCEGDMENPMRPSQORRTFTISLGEAC 720
 Db 661 QCPQPAARRRSVOLTSEKMDKVGYPKSLRKCCEGDMENPMRPSQORRTFTISLGEAC 720
 QY 721 KKFVLDCCNYITELRRQHARASHGLARSNLDEDIIAENIVSRSEFPESLWNVDELKE 780
 Db 721 KKFVLDCCNYITELRRQHARASHGLARSNLDEDIIAENIVSRSEFPESLWNVDELKE 780
 QY 781 PKKNGISTKLMNIFLKDSTITWILAVMSDKKICVADPPVTVNQOFFIDILPLPSV 840
 Db 781 PKKNGISTKLMNIFLKDSTITWILAVMSDKKICVADPPVTVNQOFFIDILPLPSV 840
 QY 841 RNEQVEIRAVLYNRYRQELKVRVELLHNPFAFCSLATTKRHHQQTITIPPKSSLSVPYVI 900
 Db 841 RNEQVEIRAVLYNRYRQELKVRVELLHNPFAFCSLATTKRHHQQTITIPPKSSLSVPYVI 900
 QY 901 VPLKTGLQVEVKAAYVHHFISDGVKRSIKVPEGIRMNKTVAVTLDPERLREGVOKE 960
 Db 901 VPLKTGLQVEVKAAYVHHFISDGVKRSIKVPEGIRMNKTVAVTLDPERLREGVOKE 960
 QY 961 DIPPADLSDOVPDTESETRILLQGTPEVAQMTEDAVDAERLKHILVTPSGCGEQNMIGWP 1020
 Db 961 DIPPADLSDOVPDTESETRILLQGTPEVAQMTEDAVDAERLKHILVTPSGCGEQNMIGWP 1020
 QY 1021 TVIAVHLDTEQWEKFGLEKRGALIELIKGYTQCLAPROPSSAFAPVGRASWTWLA 1080
 Db 1021 TVIAVHLDTEQWEKFGLEKRGALIELIKGYTQCLAPROPSSAFAPVGRASWTWLA 1080
 QY 1081 YVVKVFLAVNLIAIDSVQLCGAVKWLILILEKQKPGDQFQEDAPVIRQEMIGGLRNNNEKD 1140
 Db 1081 YVVKVFLAVNLIAIDSVQLCGAVKWLILILEKQKPGDQFQEDAPVIRQEMIGGLRNNNEKD 1140
 QY 1141 MALTAFLVLSLOAKDICEQVNSLPGSITKAGDFLEANNYNLQBSYTVATAGVLAQMG 1200
 Db 1141 MALTAFLVLSLOAKDICEQVNSLPGSITKAGDFLEANNYNLQBSYTVATAGVLAQMG 1200
 QY 1201 RLKGFPLNKFLTTAKQKRWEDSGKQLYNVEATSVALLALQLKDFDFVPPVFWRLNEQR 1260
 Db 1201 RLKGFPLNKFLTTAKQKRWEDSGKQLYNVEATSVALLALQLKDFDFVPPVFWRLNEQR 1260
 QY 1261 YVGGYGSTQATFVWFQALAQYQKDAFDHQLNLDVSLQLPSSSKIITHRIHWESASLLQ 1320

Db 1261 YVGGYGSTQATFVWFQALAQYQKDAFDHQLNLDVSLQLPSSSKIITHRIHWESASLLQ 1320
 QY 1321 SEETKENEGFTVTAAGKGGTSLVVTMYHAKAQDLTKNFKDKYTIKAPETEKRPQDA 1380
 Db 1321 SEETKENEGFTVTAAGKGGTSLVVTMYHAKAQDLTKNFKDKYTIKAPETEKRPQDA 1380
 QY 1381 KVTMLLEICTRYRGDQDATMSIIDISMTGFAFPTDLDQLANGVDVRYISKYELDKAFSD 1440
 Db 1381 KVTMLLEICTRYRGDQDATMSIIDISMTGFAFPTDLDQLANGVDVRYISKYELDKAFSD 1440
 QY 1441 RNTLIIVLDKYSHSDDECLAFKHOYENVELIQPGAVKYVAYYNLEESCTRFYHBEKSDG 1500
 Db 1441 RNTLIIVLDKYSHSDDECLAFKHOYENVELIQPGAVKYVAYYNLEESCTRFYHBEKSDG 1500
 QY 1501 KLNLKCRDELRCABENCFIQKSDDKVTLLEERLDKACEFGVDYVYKTLVKVQLSNDPDE 1560
 Db 1501 KLNLKCRDELRCABENCFIQKSDDKVTLLEERLDKACEFGVDYVYKTLVKVQLSNDPDE 1560
 QY 1561 YMALEQTIKSGSDEVQVGOORTFIPISKREALKLEBKHYLMGLSSDFWGEKPNLSY 1620
 Db 1561 YMALEQTIKSGSDEVQVGOORTFIPISKREALKLEBKHYLMGLSSDFWGEKPNLSY 1620
 QY 1621 IIGKDTWVEHWPEDBECQDBENQKQCDLGAFTESWVFGCPN 1663
 Db 1621 IIGKDTWVEHWPEDBECQDBENQKQCDLGAFTESWVFGCPN 1663

RESULT 5

AAW34616
 ID AAW34616 standard; protein; 1663 AA.

XX AAW34616;
 XX AC
 XX DT 09-APR-1998 (first entry)
 XX DE Human C3 protein mutant CV-5.
 XX KW Human; C3 protein; convertase; complement pathway protein; infection;
 KW down-regulation resistant C3 convertase; xenograft rejection; therapy;
 KW complement-mediated disease; autoimmune disease; leukaemia cell; tumour;
 KW complement-mediated response; MHC-mismatched lymphocyte; mutein.
 XX OS Homo sapiens.

Key Location/Qualifiers
 FH Misc-difference 1134
 FT Note= "R1134Q mutation"

XX WO9732981-A1.

XX 12-SEP-1997.

XX 04-MAR-1997; 97WO-GB000603.

XX 07-VAR-1996; 96GB-00004865.

XX 07-JUN-1996; 96GB-00011896.

XX 08-JUL-1996; 96GB-00014293.

XX 19-NOV-1996; 96GB-00024028.

XX (IMUT-) IMUTRAN LTD.

XX Farries TC, Harrison RA;

XX WPI; 1997-457534/42.

Modified complement pathway protein that forms C3 convertase resistant to
 down-regulation - used to exhaust the complement pathway by super-
 activation, especially for preventing graft rejection, etc.

XX Example 14; Page; 123pp; English.

XX This sequence represents a mutated human C3 protein of the invention (see

AAW34606 for wild type protein). This protein is a protein of the invention, and is a modified native complement pathway protein (A) that forms a down-regulation resistant C3 convertase. (A), their variants, fragments and conjugates are used to deplete levels of complement pathway proteins (by superactivation until one or more components are exhausted), specifically to prevent rejection of foreign material (particularly a xenograft) but also to prevent complement-mediated diseases resulting from (surgical) injury or antibody-antigen interaction in autoimmune disease, also to localise and/or amplify endogenous complement protein conversion and deposition at a specific site (e.g. a virus, infected cell or tumour), to increase sensitivity to complement-mediated responses; a particular application is eliminating any cancer cells left after surgical removal of a tumour). Also contemplated is ex vivo treatment, especially by passing blood through a matrix containing (A) (this may remove additional anaphylactic peptides and other inflammatory mediators) or killing of leukaemia cells or MHC-mismatched lymphocytes in extracted bone marrow. Since (A) is not inhibited by factor I, it can bind repeatedly to factor B (which is then inactivated), causing inactivation of the alternative pathway by consumption of factor B.

XX SQ Sequence 1663 AA;

Query Match 100.0%; Score 8605; DB 2; Length 1663;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1662; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGPTSGPSLLILLTHLPLALGSPMYSIIITPNILRLESEBETWLEAHDAGQDPVTVTVH 60
 Db 1 MGPTSGPSLLILLTHLPLALGSPMYSIIITPNILRLESEBETWLEAHDAGQDPVTVTVH 60
 QY 61 DFPKGLVLSSEKTVLTPATNMGNVFTTIPANREFKSEKGRNKFVTVQATFGTQVVEKV 120
 Db 61 DFPKGLVLSSEKTVLTPATNMGNVFTTIPANREFKSEKGRNKFVTVQATFGTQVVEKV 120
 QY 121 VLVSLSQGYLFIOFTDKTYTPGSTVLYRIETVNHKLLPVGRTVMNINPEGIPVKQDSL 180
 Db 121 VLVSLSQGYLFIOFTDKTYTPGSTVLYRIETVNHKLLPVGRTVMNINPEGIPVKQDSL 180
 QY 181 SSQQLGVLPDWDIPELVNMGMQKIRAYYENSPOQVFSTEFVEKYEIVLPSFEVIVEPTE 240
 Db 181 SSQQLGVLPDWDIPELVNMGMQKIRAYYENSPOQVFSTEFVEKYEIVLPSFEVIVEPTE 240
 QY 241 KFYIYNEKGLGVITITARELYGKVEGTAFTVIGIODGEORISLPSLXRIEDSGSEV 300
 Db 241 KFYIYNEKGLGVITITARELYGKVEGTAFTVIGIODGEORISLPSLXRIEDSGSEV 300
 QY 301 VLSRKVLDDGQNPRAEDLVGKSLYVSATVILHSGSDMVQAERSGPIVTSFYQHFTKT 360
 Db 301 VLSRKVLDDGQNPRAEDLVGKSLYVSATVILHSGSDMVQAERSGPIVTSFYQHFTKT 360
 QY 361 PKYFKGPPDLMVFNTPDGSPAYRVPVAVOGEDTVOSLTQDGVAKLSINTHPSOKPL 420
 Db 361 PKYFKGPPDLMVFNTPDGSPAYRVPVAVOGEDTVOSLTQDGVAKLSINTHPSOKPL 420
 QY 421 SITVTRKQELSEAEQATRMQALPYSTGVNSNNYLLHSLVRLTELPFGTLLNVNFLRM 480
 Db 421 SITVTRKQELSEAEQATRMQALPYSTGVNSNNYLLHSLVRLTELPFGTLLNVNFLRM 480
 QY 481 RAHEAKRYTYLTMNKRLLKAGQVREPGDLVLPISITITDFIPSLVAYTYLIGA 540
 Db 481 RAHEAKRYTYLTMNKRLLKAGQVREPGDLVLPISITITDFIPSLVAYTYLIGA 540
 QY 541 SGQREVADSVWVDKSCVGSLLVYKSGQSEDRQPVPGQMTLKIEGDHGAARVLVAVDK 600
 Db 541 SGQREVADSVWVDKSCVGSLLVYKSGQSEDRQPVPGQMTLKIEGDHGAARVLVAVDK 600
 QY 601 GFVNLKKNKLTOSKIDWVEKADIGCTPGSGKVAGVFSADGLITSSSGQGTQRAEL 660
 Db 601 GFVNLKKNKLTOSKIDWVEKADIGCTPGSGKVAGVFSADGLITSSSGQGTQRAEL 660
 QY 661 QCPQFAARRRSVLTETKMDKVGKYPKELRKCCEDGMRENPMRFSCQRRTRFISIGBAC 720
 Db 661 QCPQFAARRRSVLTETKMDKVGKYPKELRKCCEDGMRENPMRFSCQRRTRFISIGBAC 720

QY 721 KKVFLDCCNYITELRQOHARASHLGLARSNLDEDIAENIVSRSEFPESMLNVEDLUKE 780
 Db 721 KKVFLDCCNYITELRQOHARASHLGLARSNLDEDIAENIVSRSEFPESMLNVEDLUKE 780
 QY 781 PPKNGISTKLNNIFLKDSTITTWEILAVSMDSKKGICVADPPFVTVMQDFFIDLRPLYSVV 840
 Db 781 PPKNGISTKLNNIFLKDSTITTWEILAVSMDSKKGICVADPPFVTVMQDFFIDLRPLYSVV 840
 QY 841 RNQVEIRAVLYNQRONQELKVRVELLHNPACSLATTKRRHQOITITPPKSSLSVPVVI 900
 Db 841 RNQVEIRAVLYNQRONQELKVRVELLHNPACSLATTKRRHQOITITPPKSSLSVPVVI 900
 QY 901 VPLKTLQGEVEKAAVYHHFISDGVKRSKLVVPEGRMKNKTAVATLDPERLREGVQKE 960
 Db 901 VPLKTLQGEVEKAAVYHHFISDGVKRSKLVVPEGRMKNKTAVATLDPERLREGVQKE 960
 QY 961 DIPADLSOVPDTESEFTILLQGTVAQMTEDAVDAERLKHLLIUTPSCGCEQNMIGTTP 1020
 Db 961 DIPADLSOVPDTESEFTILLQGTVAQMTEDAVDAERLKHLLIUTPSCGCEQNMIGTTP 1020
 QY 1021 TVIAVHYLDBTEQWEKFGLEKQGALELIKKGYTQQLAFROPSSAFAAFVRAPESTWLTA 1080
 Db 1021 TVIAVHYLDBTEQWEKFGLEKQGALELIKKGYTQQLAFROPSSAFAAFVRAPESTWLTA 1080
 QY 1081 YVVKVPSLAVNLITIDSQVLCGAVKWLIIIEKQPGVQEDAPVTHQEMI CGLRNNNEKD 1140
 Db 1081 YVVKVPSLAVNLITIDSQVLCGAVKWLIIIEKQPGVQEDAPVTHQEMI CGLRNNNEKD 1140
 QY 1141 MALTAFLVLSQEAQDICEEVNSLPGSITKAGDLEANYMNLQRSYTVAIAGYALAQMG 1200
 Db 1141 MALTAFLVLSQEAQDICEEVNSLPGSITKAGDLEANYMNLQRSYTVAIAGYALAQMG 1200
 QY 1201 RLKGPLLNKFLTTAKDKNWEDEPGKOLYNNVATSVALIALLQLQDFDVPVVRHLNEQR 1260
 Db 1201 RLKGPLLNKFLTTAKDKNWEDEPGKOLYNNVATSVALIALLQLQDFDVPVVRHLNEQR 1260
 QY 1261 YVGGYGSTQATPMVFOALAQYQKAPDHQELNLDVSLQPSRSKITHRIHWEASLLR 1320
 Db 1261 YVGGYGSTQATPMVFOALAQYQKAPDHQELNLDVSLQPSRSKITHRIHWEASLLR 1320
 QY 1321 SEETKNEGFTVTAEGKGGTILSVVTMTHAKAKDOLTKNKDLYKTIKPAPETEKRPDA 1380
 Db 1321 SEETKNEGFTVTAEGKGGTILSVVTMTHAKAKDOLTKNKDLYKTIKPAPETEKRPDA 1380
 QY 1381 KNTMILEICTRYRGDQDATMSILDISMMTGFAPDITDQLKQLANGVDVDRYSKYELDKAFSD 1440
 Db 1381 KNTMILEICTRYRGDQDATMSILDISMMTGFAPDITDQLKQLANGVDVDRYSKYELDKAFSD 1440
 QY 1441 RNTLIIYLDKVSHEDDCLAFKHQVFNVELIOPGAVKYAYNLEESCTRYHPEKEDG 1500
 Db 1441 RNTLIIYLDKVSHEDDCLAFKHQVFNVELIOPGAVKYAYNLEESCTRYHPEKEDG 1500
 QY 1501 KLNKLCRDELRCRCAEENCFIQSDDDKVTLEERLDCACBPGVDYVYKTLVQLSNDPDE 1560
 Db 1501 KLNKLCRDELRCRCAEENCFIQSDDDKVTLEERLDCACBPGVDYVYKTLVQLSNDPDE 1560
 QY 1561 YIMAIETIKSGSDVYVQCOQRTFTSPIKCREALKLEBKHYLMWGLSDFWGKPNLSY 1620
 Db 1561 YIMAIETIKSGSDVYVQCOQRTFTSPIKCREALKLEBKHYLMWGLSDFWGKPNLSY 1620
 QY 1621 IIGKDTWHEHPDEECQDEENKQKQCDLGAFTESWVYFGCPN 1663
 Db 1621 IIGKDTWHEHPDEECQDEENKQKQCDLGAFTESWVYFGCPN 1663

RESULT 6

AAW34618

ID AAW34618 standard; protein; 1663 AA.

XX AC AAW34618;

XX AC AAW34618;

DT 09-APR-1998 (first entry)

Db 1201 RLKGLNKKFLTTAKDKNRHEDPKQLYNYEATSYALLALLQLKDFDPVPPVWNLNQR 1260
 Qy 1261 YGGGYSQTATFWFQALAQYQKADPDHQLNLDVSLQPSRSSKITHRIHWESASLLR 1320
 Db 1261 YGGGYSQTATFWFQALAQYQKADPDHQLNLDVSLQPSRSSKITHRIHWESASLLR 1320
 Qy 1321 SEETKENEFTVTAEGKGQGLTSLVWYHAKAKDQLTCKNFDLKVTIKPAPETKRPQDA 1380
 Db 1321 SEETKENEFTVTAEGKGQGLTSLVWYHAKAKDQLTCKNFDLKVTIKPAPETKRPQDA 1380
 Qy 1381 KNTWILEICRYRGDQDATWSIIDISMTGFAFDPTDQLQLANGVDVRYISKYELDKAFSD 1440
 Db 1381 KNTWILEICRYRGDQDATWSIIDISMTGFAFDPTDQLQLANGVDVRYISKYELDKAFSD 1440
 Qy 1441 RNTLIIYLDKVSHEDDCLAFKHQYFNVELIOPGAVKVVAYVYNNLEESTRFRVHPEKSDG 1500
 Db 1441 RNTLIIYLDKVSHEDDCLAFKHQYFNVELIOPGAVKVVAYVYNNLEESTRFRVHPEKSDG 1500
 Qy 1501 KLKLCRDELRCRAENCFIQKSDDKVTLBERLDKACEPGVDVYVYKTRLVKQLSNDPDE 1560
 Db 1501 KLKLCRDELRCRAENCFIQKSDDKVTLBERLDKACEPGVDVYVYKTRLVKQLSNDPDE 1560
 Qy 1561 YMAIEQTIKSGSDEVQVGQRTFISPIKCREALKLEEKKGHYLMWGLSSDFWGEKPNLSY 1620
 Db 1561 YMAIEQTIKSGSDEVQVGQRTFISPIKCREALKLEEKKGHYLMWGLSSDFWGEKPNLSY 1620
 Qy 1621 IIGKDTWVHWPEDECODEENKQCODLGAFTSMWVFCPN 1663
 Db 1621 IIGKDTWVHWPEDECODEENKQCODLGAFTSMWVFCPN 1663

RESULT 7

AAW40989
 ID AAW40989 standard; protein; 1663 AA.
 AC AAW40989;
 XX
 DT 09-APR-1998 (first entry)
 XX Human C3 protein mutant R1303X.
 DE
 KW Human; C3 protein; convertase; complement pathway protein; infection;
 KW down-regulation resistant C3 convertase; xenograft rejection; therapy;
 KW complement-mediated disease; autoimmune disease; leukaemia cell; tumour;
 KW complement-mediated response; XHC-mismatched lymphocyte; mutein.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 1303
 FT /label= Glu, Gln, Gly
 XX
 XX W09732981-A1.
 XX
 XX 12-SEP-1997.
 XX
 XX 04-MAR-1997; 97WO-GB000603.
 XX
 XX 07-MAR-1996; 96GB-00004865.
 XX 07-JUN-1996; 96GB-00011896.
 XX 08-JUL-1996; 96GB-00014293.
 XX 19-NOV-1996; 96GB-00024028.
 XX
 XX (IMUT-) IMUTRAN LTD.
 XX
 XX Farries TC, Harrison RA;
 XX
 XX WPI; 1997-457534/42.
 XX
 XX Modified complement pathway protein that forms C3 convertase resistant to
 XX down-regulation - used to exhaust the complement pathway by super-
 XX activation, especially for preventing graft rejection, etc.

PS Claim 8; Page; 123pp; English.

XX This sequence represents a mutated human C3 protein of the invention (see
 CC AAW34606 for wild type protein). This protein is a protein of the
 CC invention, and is a modified native complement pathway protein (A) that
 CC forms a down-regulation resistant C3 convertase. (A), their variants,
 CC fragments and conjugates are used to deplete levels of complement pathway
 CC proteins (by superactivation until one or more components are exhausted),
 CC specifically to prevent rejection of foreign material (particularly a
 CC xenograft) but also to prevent complement-mediated diseases resulting
 CC from (surgical) injury or antibody-antigen interaction in autoimmune
 CC disease, also to localise and/or amplify endogenous complement protein
 CC conversion and deposition at a specific site (e.g. a virus, infected cell
 CC or tumour, to increase sensitivity to complement-mediated responses; a
 CC particular application is eliminating any cancer cells left after
 CC surgical removal of a tumour). Also contemplated is ex vivo treatment,
 CC especially by passing blood through a matrix containing (A) (this may
 CC remove additional anaphylactic peptides and other inflammatory mediators)
 CC or killing of leukaemia cells or MHC-mismatched lymphocytes in extracted
 CC bone marrow. Since (A) is not inhibited by factor I, it can bind
 CC repeatedly to factor B (which is then inactivated), causing inactivation
 CC of the alternative pathway by consumption of factor B

XX Sequence 1663 AA;

Query Match 99.9%; Score 8603; DB 2; Length 1663;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1662; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGPTSGPSLLLLLLTHLPALGSPMYSIITPNILRESEETWVLEAHDAQGDVPVTVVH 60
 Db 1 MGPTSGPSLLLLLLTHLPALGSPMYSIITPNILRESEETWVLEAHDAQGDVPVTVVH 60
 Qy 61 DFFGKLVLSSEKTVLPATNMGVNTFTIPANREFKSEKGNKPVTVQATFGQVVEKV 120
 Db 61 DFFGKLVLSSEKTVLPATNMGVNTFTIPANREFKSEKGNKPVTVQATFGQVVEKV 120
 Qy 121 VLVSQSGVLFITQDKTIYTPGSTVLYRIFTVNHKLLPVGRVWVNIENPEGIPVKQDSL 180
 Db 121 VLVSQSGVLFITQDKTIYTPGSTVLYRIFTVNHKLLPVGRVWVNIENPEGIPVKQDSL 180
 Qy 181 SSQNLGVLPWDPELVNMGQWIRAYENSPQVSTFEVKEVYVLPSEFVIVEPTE 240
 Db 181 SSQNLGVLPWDPELVNMGQWIRAYENSPQVSTFEVKEVYVLPSEFVIVEPTE 240
 Qy 241 KFYIYNEXGLEVTITARELYGKVEGTAFVIFGIDGQEQRISLPESLKRIPIEDGSCEV 300
 Db 241 KFYIYNEXGLEVTITARELYGKVEGTAFVIFGIDGQEQRISLPESLKRIPIEDGSCEV 300
 Qy 301 VLSRKVLDDGVONPRAEDLVGKSLYVSATVILHSGSDMWQABERSGIPVTSFYQHFTKT 360
 Db 301 VLSRKVLDDGVONPRAEDLVGKSLYVSATVILHSGSDMWQABERSGIPVTSFYQHFTKT 360
 Qy 361 PKYFKPMFPDLWVFTNPDGSPAYRVPVAVOGEDTVQSLTQGDGVAKLSINTHPSQKPL 420
 Db 361 PKYFKPMFPDLWVFTNPDGSPAYRVPVAVOGEDTVQSLTQGDGVAKLSINTHPSQKPL 420
 Qy 421 SITVTRTKQELSEAEQATRTMQALPYSTVGNSSNNYHLVRLTELRLPGETLVNPFLLRMD 480
 Db 421 SITVTRTKQELSEAEQATRTMQALPYSTVGNSSNNYHLVRLTELRLPGETLVNPFLLRMD 480
 Qy 481 RAHEAKIRYVTVLIMNKGRLKAGQVREPQDLVPLSLITTFPIPSFRVAVYVTLIGA 540
 Db 481 RAHEAKIRYVTVLIMNKGRLKAGQVREPQDLVPLSLITTFPIPSFRVAVYVTLIGA 540
 Qy 541 SGQREVVDVSVVDVKDSCVGLSVKSGQSDRQVPCQMTLKIEGDHGARVVLVAVDK 600
 Db 541 SGQREVVDVSVVDVKDSCVGLSVKSGQSDRQVPCQMTLKIEGDHGARVVLVAVDK 600
 Qy 601 GPFVLNKKNLTKSQKIDWVEKADICCTPGSGKDYAGVFS DAGLTFITSSSQQAQRAEL 660
 Db 601 GPFVLNKKNLTKSQKIDWVEKADICCTPGSGKDYAGVFS DAGLTFITSSSQQAQRAEL 660

661 QCPQARRRRSQTLEKMDKVGKPKELKCCEDGMENPMFSCQRTFRISLGEAC 720
 661 QCPQARRRRSQTLEKMDKVGKPKELKCCEDGMENPMFSCQRTFRISLGEAC 720
 721 KVFLOCCNYITELRQHARASHGLARNSLDEDEIIAENIVSRSEPPESWLNVEDLKE 780
 721 KVFLOCCNYITELRQHARASHGLARNSLDEDEIIAENIVSRSEPPESWLNVEDLKE 780
 781 PPKXGISTKLMIIFIKDSITTWELAVNSDKKICVADPEFTVMQDFIDRLPVSIV 840
 781 PPKXGISTKLMIIFIKDSITTWELAVNSDKKICVADPEFTVMQDFIDRLPVSIV 840
 841 RNEQVIRAVLYNRYNQELKVRVLLHNPAFCSLATTKRRHQOTIIPPKSSLSVPYVI 900
 841 RNEQVIRAVLYNRYNQELKVRVLLHNPAFCSLATTKRRHQOTIIPPKSSLSVPYVI 900
 901 VPLKTGLQEVKAAVYHHFISDGVKSLKVVPEGIRNKNKTAVRTLDPERLGRGVQKE 960
 901 VPLKTGLQEVKAAVYHHFISDGVKSLKVVPEGIRNKNKTAVRTLDPERLGRGVQKE 960
 961 DIPPADLSQVPTSETRIILQGTVPVQMTEDAVDAERLKHLLIVTPSGCGEQNMIGWP 1020
 961 DIPPADLSQVPTSETRIILQGTVPVQMTEDAVDAERLKHLLIVTPSGCGEQNMIGWP 1020
 1021 TVIAVHYLDETEQWEGFLEKKGALIKKGYTQCLAFROPSSAFAPVFRAPSTWLTA 1080
 1021 TVIAVHYLDETEQWEGFLEKKGALIKKGYTQCLAFROPSSAFAPVFRAPSTWLTA 1080
 1081 YVVKVFLANLAIQSVLCAVAKWLLILEKQKPGVFEQDAPVHQEMIGLGNNEKD 1140
 1081 YVVKVFLANLAIQSVLCAVAKWLLILEKQKPGVFEQDAPVHQEMIGLGNNEKD 1140
 1141 MALTAFLVLSQPAKIDCEQVNSLPGSITKAGDFLEANYMNLQSYTVAAGVALAQMG 1200
 1141 MALTAFLVLSQPAKIDCEQVNSLPGSITKAGDFLEANYMNLQSYTVAAGVALAQMG 1200
 1201 RLKGPLNKLFTTAKDKNWEPPGKQVNVETSVALLALQLKDFDPVPPVWVWLEQ 1260
 1201 RLKGPLNKLFTTAKDKNWEPPGKQVNVETSVALLALQLKDFDPVPPVWVWLEQ 1260
 1261 YCGGVGSGTOATPMVFPALAAQKADPHQELNLDVSLQLPSRSKITHRIHWSASLLR 1320
 1261 YCGGVGSGTOATPMVFPALAAQKADPHQELNLDVSLQLPSRSKITHRIHWSASLLR 1320
 1321 SEETKNEGFTVTAEGKGQGTLSVVTMYHAKAKDQJTCNKKDLKVTIKPAPETEKRPQDA 1380
 1321 SEETKNEGFTVTAEGKGQGTLSVVTMYHAKAKDQJTCNKKDLKVTIKPAPETEKRPQDA 1380
 1381 KNTWILEICTRYRGDQATMSILDTSMTGFPAPDITDQLKQANGVDVRSIKYELDKAFSD 1440
 1381 KNTWILEICTRYRGDQATMSILDTSMTGFPAPDITDQLKQANGVDVRSIKYELDKAFSD 1440
 1441 RNTLIYLDKVSSEDDCLAFKHQVFNVELTOPGAKVYAYVYNNLEESCTFRYPHEKEDG 1500
 1441 RNTLIYLDKVSSEDDCLAFKHQVFNVELTOPGAKVYAYVYNNLEESCTFRYPHEKEDG 1500
 1501 KLNKLCRDELRCRCAENCFIQKSDDKVTLEBLDKACBPQVDVYVTKLVQVLSNDFE 1560
 1501 KLNKLCRDELRCRCAENCFIQKSDDKVTLEBLDKACBPQVDVYVTKLVQVLSNDFE 1560
 1561 YMAIEQTIKSGDVEVQVQGTFTSPIKCRALKLEEKHKLWGLSSDFWGEKPNLSY 1620
 1561 YMAIEQTIKSGDVEVQVQGTFTSPIKCRALKLEEKHKLWGLSSDFWGEKPNLSY 1620
 1621 IIGKDTWVHWPBDEDCQENKQKQODLGAFTESWVFGCPN 1663
 1621 IIGKDTWVHWPBDEDCQENKQKQODLGAFTESWVFGCPN 1663

RESULT 8
 AAW34608
 ID AAW34608 standard; protein; 1663 AA.
 XX

AC AAW34608;
 XX 09-APR-1998 (first entry)
 XX Human C3 protein mutant C3M-26.
 XX Human; C3 protein; convertase; complement pathway protein; infection;
 XX down-regulation resistant C3 convertase; xenograft rejection; therapy;
 XX complement-mediated disease; autoimmune disease; leukaemia cell; tumour;
 XX complement-mediated response; MHC-mismatched lymphocyte; mutain.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 XX Misc-difference 1303 /note= "Arg to Gln mutation"
 XX Misc-difference 1320 /note= "Arg to Gln mutation"
 XX WO9732981-A1.
 XX 12-SEP-1997.
 XX 04-MAR-1997; 97WO-GB0006503.
 XX 07-MAR-1996; 96GB-00004855.
 XX 07-JUN-1996; 96GB-00011896.
 XX 08-JUL-1996; 96GB-00014293.
 XX 19-NOV-1996; 96GB-00024028.
 XX (IMUT-) IMUTRAN LTD.
 XX Farries TC, Harrison RA;
 XX WPI; 1997-457534/42.
 XX Modified complement pathway protein that forms C3 convertase resistant to
 XX down-regulation - used to exhaust the complement pathway by super-
 XX activation, especially for preventing graft rejection, etc.
 XX Example 4; Page; 123pp; English.
 XX This sequence represents a mutated human C3 protein of the invention (see
 XX AAW34608 for wild type protein). This protein is a protein of the
 XX invention, and is a modified native complement pathway protein (A) that
 XX forms a down-regulation resistant C3 convertase. (A), their variants,
 XX fragments and conjugates are used to deplete levels of complement pathway
 XX proteins (by superactivation until one or more components are exhausted),
 XX specifically to prevent rejection of foreign material (particularly a
 XX xenograft) but also to prevent complement-mediated diseases resulting
 XX from (surgical) injury or antibody-antigen interaction in autoimmune
 XX disease, also to localise and/or amplify endogenous complement protein
 XX conversion and deposition at a specific site (e.g. a virus, infected cell
 XX or tumour, to increase sensitivity to complement-mediated responses; a
 XX particular application is eliminating any cancer cells left after
 XX surgical removal of a tumour). Also contemplated is ex vivo treatment,
 XX especially by passing blood through a matrix containing (a) (this may
 XX remove additional anaphylactic peptides and other inflammatory mediators)
 XX or killing of leukaemia cells or MHC-mismatched lymphocytes in extracted
 XX bone marrow. Since (A) is not inhibited by factor I, it can bind
 XX repeatedly to factor B (which is then inactivated), causing inactivation
 XX of the alternative pathway by consumption of factor B
 XX Sequence 1663 AA;
 XX Query Match 99.9%; Score 8601; DB 2; Length 1663;
 XX Best Local Similarity 99.9%; Pred. No. 0;
 XX Matches 1661; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGPTSGPSLLLLLTHLPALGSPMYSIITPNILRLESEETWVLEAHDAGQGVPTVTVH 60
 Db 1 MGPTSGPSLLLLLTHLPALGSPMYSIITPNILRLESEETWVLEAHDAGQGVPTVTVH 60

QY 61 DPGKGLVLSSEKTLTPATNMGNTFTIPANREKSKGRNKFVTVQATFGTQVSEKV 120
DB 61 DPGKGLVLSSEKTLTPATNMGNTFTIPANREKSKGRNKFVTVQATFGTQVSEKV 120
QY 121 VLVSQSGLYFIQTDKTYITPGSTVLYRIETVNHKLLPVGRITVMNIENPEGIPVKQDSL 180
DB 121 VLVSQSGLYFIQTDKTYITPGSTVLYRIETVNHKLLPVGRITVMNIENPEGIPVKQDSL 180
QY 181 SSQNLGVLPLSWDIPELVNMGMKIRAYYENSPQOVSTEFPEVEKYLPLSPSEVIVEPTE 240
DB 181 SSQNLGVLPLSWDIPELVNMGMKIRAYYENSPQOVSTEFPEVEKYLPLSPSEVIVEPTE 240
QY 241 KFYIYNEKGLLEVITITAFILYKKGVEGTAFVIFGIDGQORISLPSLKRIFIDGSGEV 300
DB 241 KFYIYNEKGLLEVITITAFILYKKGVEGTAFVIFGIDGQORISLPSLKRIFIDGSGEV 300
QY 301 VLSRKVLIDGQNPRAEDLVGKSLYVSATVILHSGSDMWQAERSGIPITSPYQHFTKT 360
DB 301 VLSRKVLIDGQNPRAEDLVGKSLYVSATVILHSGSDMWQAERSGIPITSPYQHFTKT 360
QY 361 PKYFKPGNPDLMVFTNPDGSPAYRYPVAVQGEDTVQSLTQDGVAKLSINTHPSOKPL 420
DB 361 PKYFKPGNPDLMVFTNPDGSPAYRYPVAVQGEDTVQSLTQDGVAKLSINTHPSOKPL 420
QY 421 SITVTRTKQELSEAEQATRTMQALPYSTVGNNSNNYLHLSVLRTLPFGTLLNVNELLRMD 480
DB 421 SITVTRTKQELSEAEQATRTMQALPYSTVGNNSNNYLHLSVLRTLPFGTLLNVNELLRMD 480
QY 481 RAHAERKRYTYLLIMNKRLLKAGQVREPQGLVLPISITTDTPSPRLVAVYTLGA 540
DB 481 RAHAERKRYTYLLIMNKRLLKAGQVREPQGLVLPISITTDTPSPRLVAVYTLGA 540
QY 541 SQREWVADSVVMDKQSCVGLSVKSGQSEDRQVPVGGQMTLKIEGDHGAHVILVAVDK 600
DB 541 SQREWVADSVVMDKQSCVGLSVKSGQSEDRQVPVGGQMTLKIEGDHGAHVILVAVDK 600
QY 601 GFVLNKNKLTQSKIWDVVEKADIGCTPGSGDYAGVPSDAGLTPTSSSGQGTQRAEL 660
DB 601 GFVLNKNKLTQSKIWDVVEKADIGCTPGSGDYAGVPSDAGLTPTSSSGQGTQRAEL 660
QY 661 QCQPAARRRSVQLTEKMDKVGKYPKELKCECDGMRENPMRPFSCQRTTIFISLGEAC 720
DB 661 QCQPAARRRSVQLTEKMDKVGKYPKELKCECDGMRENPMRPFSCQRTTIFISLGEAC 720
QY 721 KKVFLDCCNVITELRQARASHLGLARSLNLDIEDITAEENIVSRSEFPESLWNVEDLKE 780
DB 721 KKVFLDCCNVITELRQARASHLGLARSLNLDIEDITAEENIVSRSEFPESLWNVEDLKE 780
QY 781 PPQNGISTKLMMIFLKOSITWEILAVMSDKKGIQVADPFVTVMQDFIDRLPYSVV 840
DB 781 PPQNGISTKLMMIFLKOSITWEILAVMSDKKGIQVADPFVTVMQDFIDRLPYSVV 840
QY 841 RNEQVEIRAVLYNQELKVRVVELLHNPFAFCSLATTKRHOQITTIIPKSSLSVPVVI 900
DB 841 RNEQVEIRAVLYNQELKVRVVELLHNPFAFCSLATTKRHOQITTIIPKSSLSVPVVI 900
QY 901 VPLKTLGQVEVKAAYVHHFTSDGVRSLKVVPEGIRMNKTAVRDLDERLREGVQKE 960
DB 901 VPLKTLGQVEVKAAYVHHFTSDGVRSLKVVPEGIRMNKTAVRDLDERLREGVQKE 960
QY 961 DIPPADLSQVDPDTESETRILLQGTTPVAQMTEDAVDAERLKLHIVTPSGCGEQNMIGTTP 1020
DB 961 DIPPADLSQVDPDTESETRILLQGTTPVAQMTEDAVDAERLKLHIVTPSGCGEQNMIGTTP 1020
QY 1021 TVIANVHLDTEQNEKFLKRGQALELTKGVTQGLAPQSSAAFAVKAPSTWTLTA 1080
DB 1021 TVIANVHLDTEQNEKFLKRGQALELTKGVTQGLAPQSSAAFAVKAPSTWTLTA 1080
QY 1081 YVVRVFLAVNLIAIDSVLQCGAVKWLILEKQKPDGVQFQEDAPVHCEMIGGLRNNNEKD 1140
DB 1081 YVVRVFLAVNLIAIDSVLQCGAVKWLILEKQKPDGVQFQEDAPVHCEMIGGLRNNNEKD 1140
QY 1141 MALTAFLVLSIQEAKDICEEQVNSLPGSITKAGDFLEANNYLQRSYTVAIAGYALQMG 1200

DB 1141 MALTAFLVLSIQEAKDICEEQVNSLPGSITKAGDFLEANNYLQRSYTVAIAGYALQMG 1200
QY 1201 RLKGPLLNKFLTTAKDKRWEDPGKQLYNVEATSYALLAILQLKDFDPVPVVRWLNQOR 1260
DB 1201 RLKGPLLNKFLTTAKDKRWEDPGKQLYNVEATSYALLAILQLKDFDPVPVVRWLNQOR 1260
QY 1261 YYGCGYGSTQATFMVFQALAQYQKDAPDHQLNLDVSLQPLSPSSKITHRIHWESASLLR 1320
DB 1261 YYGCGYGSTQATFMVFQALAQYQKDAPDHQLNLDVSLQPLSPSSKITHRIHWESASLLR 1320
QY 1321 SEETENEGFTVTAEGKGGTILSVVMYHAKAKQDLTCNKFDLKVTIKPAPETEKPODA 1380
DB 1321 SEETENEGFTVTAEGKGGTILSVVMYHAKAKQDLTCNKFDLKVTIKPAPETEKPODA 1380
QY 1381 KNTMLEICTRYRGQDATMSILDISMMTGFAPPTDDLKOLANGVDRYISKYELDKAFSD 1440
DB 1381 KNTMLEICTRYRGQDATMSILDISMMTGFAPPTDDLKOLANGVDRYISKYELDKAFSD 1440
QY 1441 RNTLIYLDKVSHSEDDCLAFKHQYFNVELIQFGAVKYAYYNLBESCTRFVHPEKEDG 1500
DB 1441 RNTLIYLDKVSHSEDDCLAFKHQYFNVELIQFGAVKYAYYNLBESCTRFVHPEKEDG 1500
QY 1501 KLNKLCRDELCLCAEENCFIOKSDDKVTLBERLKACEPGVDVYVYKTRLVKVLSDNDFDE 1560
DB 1501 KLNKLCRDELCLCAEENCFIOKSDDKVTLBERLKACEPGVDVYVYKTRLVKVLSDNDFDE 1560
QY 1561 YVMAIEOTIKSGSDVQVQOQRTFISPKCREALKLEKHYLMWGLSSDFWGEKPNLSY 1620
DB 1561 YVMAIEOTIKSGSDVQVQOQRTFISPKCREALKLEKHYLMWGLSSDFWGEKPNLSY 1620
QY 1621 IIGKUTVHWHPEDEECQDEENQKOCQDLGAFTESMVVFQCPN 1663
DB 1621 IIGKUTVHWHPEDEECQDEENQKOCQDLGAFTESMVVFQCPN 1663

RESULT 9

AAR94028

ID AAR94028 standard; protein; 1663 AA.

AC AAR94028;

DT 21-MAY-1996 (first entry)

XX Human C3 precursor.

XX C3 protein; convertase; Factor I; Factor H; complement.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Peptide 1..22

XX FT /label= Sig_peptide

XX FT 23..667

XX FT /note= "C3 beta chain"

XX FT 668..671

XX FT /note= "amino acids 668-671 are removed when the

XX FT precursor is cleaved into the alpha and beta chains"

XX FT 672..1663

XX FT /note= "C3 alpha chain"

XX WO9607738-A2.

XX 14-MAR-1996.

XX 08-SEP-1995; 95WO-CB002121.

XX 08-SEP-1994; 94GB-00018147.

XX 04-MAY-1995; 95GB-00009102.

XX (IMUT-) IMUTRAN LTD.

XX Harrison RA, Farries TC;

PI

XX WPI; 1996-171613/17.
 DR N-PSDB; AAT17738.
 XX
 PT Mutant complement pathway protein forming stable C3 convertase - for
 PT generalised complement depletion or localised complement activation.
 XX
 PS Disclosure; Fig 1; 81pp; English.
 XX
 CC Human C3 protein (AAR94028) was produced by expression of a cDNA sequence
 CC (AAT17738) isolated from a human liver cDNA library. C3 is a complement
 CC pathway protein that is susceptible to cleavage by Factor I and is also
 CC susceptible to the inhibitory action of Factor H. Mutants of C3 (AAR94029
 CC and AAR94030) have been produced by site-directed mutagenesis. These
 CC mutants can be used to super-activate the complement system, or to induce
 CC localised super-activation at a specific target to increase the target's
 CC sensitivity to complement-mediated destruction
 XX
 XX Sequence 1663 AA;
 SQ

Query Match 99.9%; Score 8600; DB 2; Length 1663;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1662; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGPTSGSLLLLLTHPLALGSPMYSIIITPNILRESEETWVLEAHDAQDQVPTVTVH 60
 DB 1 MGPTSGSLLLLLTHPLALGSPMYSIIITPNILRESEETWVLEAHDAQDQVPTVTVH 60
 QY 61 DFPCKVLVSEKTVLTPATNHMGNTFTIPANREFKSEKGNKFNFTVQATFGVVEKV 120
 DB 61 DFPCKVLVSEKTVLTPATNHMGNTFTIPANREFKSEKGNKFNFTVQATFGVVEKV 120
 QY 121 VLVSQSGYLFIOTDKTIVTPGTVLYRFTVNHKLLPVGRVWVNIENPESGIPVKQDSL 180
 DB 121 VLVSQSGYLFIOTDKTIVTPGTVLYRFTVNHKLLPVGRVWVNIENPESGIPVKQDSL 180
 QY 181 SSQNLGVPLSWDIFELVNMGWKIRAYENSPQVSTFTEVEKVEYVLPSEFVIVPTE 240
 DB 181 SSQNLGVPLSWDIFELVNMGWKIRAYENSPQVSTFTEVEKVEYVLPSEFVIVPTE 240
 QY 241 KFYIYNEKGLVITITARFLYGKVEGTAFVIFGIDQGEQRISLPESLKRIPIEDGSGEV 300
 DB 241 KFYIYNEKGLVITITARFLYGKVEGTAFVIFGIDQGEQRISLPESLKRIPIEDGSGEV 300
 QY 301 VLSKRVLDGVONPRADLVGKSLYVSATVILHSGSDMVQARSGIPITVSPYQIHFTKT 360
 DB 301 VLSKRVLDGVONPRADLVGKSLYVSATVILHSGSDMVQARSGIPITVSPYQIHFTKT 360
 QY 361 PKYFKPGMPDLVFNTPDGSFAYRVPVAVQGEDTVQSLTQGDGVAKLSINTHPSQKPL 420
 DB 361 PKYFKPGMPDLVFNTPDGSFAYRVPVAVQGEDTVQSLTQGDGVAKLSINTHPSQKPL 420
 QY 421 SITVTRKKSELSAEQATRMQALPYSTVGNSSNVLHLSVLTRELPGSTLWNNFLRMD 480
 DB 421 SITVTRKKSELSAEQATRMQALPYSTVGNSSNVLHLSVLTRELPGSTLWNNFLRMD 480
 QY 481 RAHEAKIRYTYILIMNKGRLLAGROVREPQGDVLVPLSITTFIPFRLVAYTYLIGA 540
 DB 481 RAHEAKIRYTYILIMNKGRLLAGROVREPQGDVLVPLSITTFIPFRLVAYTYLIGA 540
 QY 541 SGQREVVDSDVVDKSDCVGSLVKSQSGSEDRQVPVQGMQLKIEGDHGARVVLVAVDK 600
 DB 541 SGQREVVDSDVVDKSDCVGSLVKSQSGSEDRQVPVQGMQLKIEGDHGARVVLVAVDK 600
 QY 601 GVFLNKKNLTKGKTDWVEKADIGCTFGSGKQVAGVFSAGLTFPTSSSQQTQAQRAEL 660
 DB 601 GVFLNKKNLTKGKTDWVEKADIGCTFGSGKQVAGVFSAGLTFPTSSSQQTQAQRAEL 660
 QY 661 QCPOPAARRRRSQTLEKMDKVGKPKELRKCCEGDMRENPMRPFSCQRRTRFISLGEAC 720
 DB 661 QCPOPAARRRRSQTLEKMDKVGKPKELRKCCEGDMRENPMRPFSCQRRTRFISLGEAC 720
 QY 721 KKVFLDCCNITELRRQHARASHGLARSNLDEDIABENIVSRSEFFESLMWVEDLKE 780

DB 721 KKVFLDCCNITELRRQHARASHGLARSNLDEDIABENIVSRSEFFESLMWVEDLKE 780
 QY 781 PPKNGISTKLMNIFLKDSITTEILAVSMDKKGICVADPPFVTVMQDFIDLRLPYSV 840
 DB 781 PPKNGISTKLMNIFLKDSITTEILAVSMDKKGICVADPPFVTVMQDFIDLRLPYSV 840
 QY 841 RNEQVEIRAVLYNVRQNOELKURVELLHNPASCSLATTKRHQOITITPPKSSISVPVI 900
 DB 841 RNEQVEIRAVLYNVRQNOELKURVELLHNPASCSLATTKRHQOITITPPKSSISVPVI 900
 QY 901 VPLTKGLOEVEVAAVYHHFISDGVKSLKVVPEGIRMNKTVAVRTLPDLREGVQKE 960
 DB 901 VPLTKGLOEVEVAAVYHHFISDGVKSLKVVPEGIRMNKTVAVRTLPDLREGVQKE 960
 QY 961 DIPADLSQVPTDESTRILLQSTTPVAQMTEDAVDAERLKLIVTPSGCGEQNMIGWP 1020
 DB 961 DIPADLSQVPTDESTRILLQSTTPVAQMTEDAVDAERLKLIVTPSGCGEQNMIGWP 1020
 QY 1021 TVIAVHYLDETEQWEXFLEKRGALALELIKGYTQOLAFPROPSAFAAFVIRAPSTWLT 1080
 DB 1021 TVIAVHYLDETEQWEXFLEKRGALALELIKGYTQOLAFPROPSAFAAFVIRAPSTWLT 1080
 QY 1081 YVWVFSLANVLLAIDSQLCGAVKWLIILEKOKPGVFOEDAPVHOEMIGLNNNEKD 1140
 DB 1081 YVWVFSLANVLLAIDSQLCGAVKWLIILEKOKPGVFOEDAPVHOEMIGLNNNEKD 1140
 QY 1141 MALTAFLVILSIOBAKOICEBOVNSLPGSTIKAGDFLEANYMNLORSYVAIAGYALQMG 1200
 DB 1141 MALTAFLVILSIOBAKOICEBOVNSLPGSTIKAGDFLEANYMNLORSYVAIAGYALQMG 1200
 QY 1201 RLKGPLLNKFLTTAKOKNRWEDPGKOLYNVEATSVALLALLQKDFDVPVFWLWNEQ 1260
 DB 1201 RLKGPLLNKFLTTAKOKNRWEDPGKOLYNVEATSVALLALLQKDFDVPVFWLWNEQ 1260
 QY 1261 YGGGYGSTQATFWFOALAQYOKDAPDHOELNLDVSLPSSSKI THRIHWSASLLR 1320
 DB 1261 YGGGYGSTQATFWFOALAQYOKDAPDHOELNLDVSLPSSSKI THRIHWSASLLR 1320
 QY 1321 SEETKENEGTVAEKGQGTLSVVTWYHAKAKDQLTCKNFDLKVTIKPAPETEKRPQDA 1380
 DB 1321 SEETKENEGTVAEKGQGTLSVVTWYHAKAKDQLTCKNFDLKVTIKPAPETEKRPQDA 1380
 QY 1381 KNTMILICTRYGDODATMSILDISMTGFPADTDDLKOLANGVDVYISKVELDKAFSD 1440
 DB 1381 KNTMILICTRYGDODATMSILDISMTGFPADTDDLKOLANGVDVYISKVELDKAFSD 1440
 QY 1441 RNTLIIYLDKVSHEDDCLAFKHQYFNVELIOPGAVKYVAYYNLEESCTFYHPEKEDG 1500
 DB 1441 RNTLIIYLDKVSHEDDCLAFKHQYFNVELIOPGAVKYVAYYNLEESCTFYHPEKEDG 1500
 QY 1501 KLNLKCEDELRCRABENCFIQKSDDDKVTLEERLDKACPGVDVYKTLVKVQSLNDPDE 1560
 DB 1501 KLNLKCEDELRCRABENCFIQKSDDDKVTLEERLDKACPGVDVYKTLVKVQSLNDPDE 1560
 QY 1561 YMAIEOTIKSGSEVQVQGGRTFISPIKCRBALKEKKHYLMWGLSSDFWGBKPNLSY 1620
 DB 1561 YMAIEOTIKSGSEVQVQGGRTFISPIKCRBALKEKKHYLMWGLSSDFWGBKPNLSY 1620
 QY 1621 IIGKDTVWEHWPBEDECOBENKQCODLGAFTESWVFGCPN 1663
 DB 1621 IIGKDTVWEHWPBEDECOBENKQCODLGAFTESWVFGCPN 1663

RESULT 10
 AAW34627
 ID AAW34627 standard; protein; 1663 AA.
 XX
 AC AAW34627;
 XX
 DT 09-APR-1998 (first entry)
 XX
 DE Human C3 protein mutant FT-5.

XX Human; C3 protein; convertase; complement pathway protein; infection;
 KW down-regulation resistant C3 convertase; xenograft rejection; therapy;
 KW complement-mediated disease; autoimmune disease; leukaemia cell; tumour;
 KW complement-mediated response; MHC-mismatched lymphocyte; mutein.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT Misc-difference 1661 /note= "C1661S mutation"
 XX

XX WO9732981-A1.

XX 12-SEP-1997.

XX 04-MAR-1997; 97WO-GB000603.

XX 07-MAR-1996; 96GB-00004865.

XX 07-JUN-1996; 96GB-00011896.

XX 08-JUL-1996; 96GB-00014293.

XX 19-NOV-1996; 96GB-00024028.

XX (IMUT-) IMUTRAN LTD.

XX Farries TC, Harrison RA;

XX WPI; 1997-457534/42.

XX Modified complement pathway protein that forms C3 convertase resistant to

XX down-regulation - used to exhaust the complement pathway by super-

XX activation, especially for preventing graft rejection, etc.

XX Example 17; Page; 123pp; English.

XX This sequence represents a mutated human C3 protein of the invention (see
 CC AW3466 for wild type protein). This protein is a protein of the
 CC invention, and is a modified native complement pathway protein (A) that
 CC forms a down-regulation resistant C3 convertase. (A), their variants,
 CC fragments and conjugates are used to deplete levels of complement pathway
 CC proteins (by superactivation until one or more components are exhausted),
 CC specifically to prevent rejection of foreign material (particularly a
 CC xenograft) but also to prevent complement-mediated diseases resulting
 CC from (surgical) injury or antibody-antigen interaction in autoimmune
 CC diseases, also to localise and/or amplify endogenous complement protein
 CC conversion and deposition at a specific site (e.g. a virus, infected cell
 CC or tumour, to increase sensitivity to complement-mediated responses; a
 CC particular application is eliminating any cancer cells left after
 CC surgical removal of a tumour). Also contemplated is ex vivo treatment,
 CC especially by passing blood through a matrix containing (A) (this may,
 CC remove additional anaphylactic peptides and other inflammatory mediators)
 CC or killing of leukaemia cells or MHC-mismatched lymphocytes in extracted
 CC bone marrow. Since (A) is not inhibited by factor I, it can bind
 CC repeatedly to factor B (which is then inactivated), causing inactivation
 CC of the alternative pathway by consumption of factor B

XX Sequence 1663 AA;

Query Match 99.98; Score 8599; DB 2; Length 1663;
 Best Local Similarity 99.98; Pred. No. 0;
 Matches 1662; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGPTSGPSLLLLTHPLALGSPFMSIITPNILRLESEETVLEAHDAQGDPVTVTVH 60

DB 1 MGPTSGPSLLLLTHPLALGSPFMSIITPNILRLESEETVLEAHDAQGDPVTVTVH 60

QY 61 DFGCKLVLSSEKTVLPATNMGNTFTFIPANREFKSEKGNKFTVQATFGTOVKEV 120

DB 61 DFGCKLVLSSEKTVLPATNMGNTFTFIPANREFKSEKGNKFTVQATFGTOVKEV 120

QY 121 VLVSQGYLFQTDKTIYTPGSTVLYRIFTVNHKLLPVGRTVMNENPEGIPVKQDSL 180

DB 121 VLVSQGYLFQTDKTIYTPGSTVLYRIFTVNHKLLPVGRTVMNENPEGIPVKQDSL 180

QY 181 SSQNQLGVPLSWDIPELVNMQWKIRAYENSPOQVSTEFEBKVKYVLPSEFVIVPEPTE 240
 DB 181 SSQNQLGVPLSWDIPELVNMQWKIRAYENSPOQVSTEFEBKVKYVLPSEFVIVPEPTE 240
 QY 241 KFYIYNKGLVITITARELYGKVEGTAFVIFGIDGEGEQRISLPESLKRIPIEDGSGEV 300
 DB 241 KFYIYNKGLVITITARELYGKVEGTAFVIFGIDGEGEQRISLPESLKRIPIEDGSGEV 300
 QY 301 VLSRKVLDDGVONPRAEDLVGKSLVSVATVILHSGSDMVQAFERSGIPVITSPYQIHFTKT 360
 DB 301 VLSRKVLDDGVONPRAEDLVGKSLVSVATVILHSGSDMVQAFERSGIPVITSPYQIHFTKT 360
 QY 361 PKYFKEGMPFDLMVFTNPDGSPAYRVPVAVQGEDTVQSLTQGGDGVAKLSINTHPSOKPL 420
 DB 361 PKYFKEGMPFDLMVFTNPDGSPAYRVPVAVQGEDTVQSLTQGGDGVAKLSINTHPSOKPL 420
 QY 421 SITVTRTKQELSEAFQATRTQALPYSTVGNNSNVLHLSVLRTELRCGETLVNWFLLRMD 480
 DB 421 SITVTRTKQELSEAFQATRTQALPYSTVGNNSNVLHLSVLRTELRCGETLVNWFLLRMD 480
 QY 481 RAHEAKIRYTYTLIMNKGRLKAGQVREPQGDLYVLPFSLTITDFIPSPRLVAYTYTLIGA 540
 DB 481 RAHEAKIRYTYTLIMNKGRLKAGQVREPQGDLYVLPFSLTITDFIPSPRLVAYTYTLIGA 540
 QY 541 SGQREVVADSVMDVKDCVGLVVKSGSDSDROPVPGQOMTKIEGDHGRVVLVAVDK 600
 DB 541 SGQREVVADSVMDVKDCVGLVVKSGSDSDROPVPGQOMTKIEGDHGRVVLVAVDK 600
 QY 601 GFVLNKNKLTQSKIMDVVERADIGCTPGSGKYAGVFS DAGLTFTSSSQOQTAQRAEL 660
 DB 601 GFVLNKNKLTQSKIMDVVERADIGCTPGSGKYAGVFS DAGLTFTSSSQOQTAQRAEL 660
 QY 661 QCPQPAARRRSVOLTEKMDKVGYPKELKCCEDGMRNENPMBSPCORRTFISLGEAC 720
 DB 661 QCPQPAARRRSVOLTEKMDKVGYPKELKCCEDGMRNENPMBSPCORRTFISLGEAC 720
 QY 721 KKVFLDCCNYITELRQRHARASHLGLARSNLDEDIIAENIVSRSEFPESWLMVVDLKE 780
 DB 721 KKVFLDCCNYITELRQRHARASHLGLARSNLDEDIIAENIVSRSEFPESWLMVVDLKE 780
 QY 781 PPNGISTKLMLIFUKDSTITWEILLAVMSDKKGCIVADPEVTVMQDFFIDLRLPSVW 840
 DB 781 PPNGISTKLMLIFUKDSTITWEILLAVMSDKKGCIVADPEVTVMQDFFIDLRLPSVW 840
 QY 841 RNEQVEIRAVLYNRQNELKVRVLLHNPAFCSLATTKRHHQQTITIPPXSSLSVPYVI 900
 DB 841 RNEQVEIRAVLYNRQNELKVRVLLHNPAFCSLATTKRHHQQTITIPPXSSLSVPYVI 900
 QY 901 VPLKTLQGEVEVKAAVYHHFISDGVKSLKVVPGEIRMNKTAVVRTLDPERLGRGVQKE 960
 DB 901 VPLKTLQGEVEVKAAVYHHFISDGVKSLKVVPGEIRMNKTAVVRTLDPERLGRGVQKE 960
 QY 961 DIPPADLSQVDPDTESETRILLQGTPTVAQMTEDAVDAERLKHLLIVTPSGCGEQNMIGTP 1020
 DB 961 DIPPADLSQVDPDTESETRILLQGTPTVAQMTEDAVDAERLKHLLIVTPSGCGEQNMIGTP 1020
 QY 1021 TVIAVHVLDETEQWEKFGLEKQGALELIIKGYTQQLAFROPSSAFARAPSTWLTA 1080
 DB 1021 TVIAVHVLDETEQWEKFGLEKQGALELIIKGYTQQLAFROPSSAFARAPSTWLTA 1080
 QY 1081 YVVKFVSLAVNLIAIDSQVLGAVKWLILEKQKPDGVFQEDAPVHQEMIGLNNNEKD 1140
 DB 1081 YVVKFVSLAVNLIAIDSQVLGAVKWLILEKQKPDGVFQEDAPVHQEMIGLNNNEKD 1140
 QY 1141 MALTAFLVLSLOEAKDICEEYVNSLPGSITKAGDFLEANYMNLORSYTVATAGYALQMG 1200
 DB 1141 MALTAFLVLSLOEAKDICEEYVNSLPGSITKAGDFLEANYMNLORSYTVATAGYALQMG 1200
 QY 1201 RLKGPLLNKFLTTAKDKRWEDPGKQLYNVEATSYALLALLQLKDFDVPVPPVWMLNEQR 1260
 DB 1201 RLKGPLLNKFLTTAKDKRWEDPGKQLYNVEATSYALLALLQLKDFDVPVPPVWMLNEQR 1260

QY 1261 YGGGYSQTATFWMVFOALAQYQKADPDHOLNLDVSLQPSRSSKITHRIHWESASLLR 1320
 Db 1261 YGGGYSQTATFWMVFOALAQYQKADPDHOLNLDVSLQPSRSSKITHRIHWESASLLR 1320
 QY 1321 SEETKNEGTVTAEGKGQOTLSVWYMYHAKAKDQLTCKNPKLKVTKIPAPETEKRPDQA 1380
 Db 1321 SEETKNEGTVTAEGKGQOTLSVWYMYHAKAKDQLTCKNPKLKVTKIPAPETEKRPDQA 1380
 QY 1381 KNTMILBICTRYRGDQDATMSIILDSMTGFAFPTDLDKOLANGVDVRIKSYELDKAFSD 1440
 Db 1381 KNTMILBICTRYRGDQDATMSIILDSMTGFAFPTDLDKOLANGVDVRIKSYELDKAFSD 1440
 QY 1441 RNTLIIYLDKVSHEDECLAFKHQYFNVELLPQGAVKYAYYNLEESCTFYHPEKEDG 1500
 Db 1441 RNTLIIYLDKVSHEDECLAFKHQYFNVELLPQGAVKYAYYNLEESCTFYHPEKEDG 1500
 QY 1501 KLNKLCRDELCRAEENCFIQKSDDKVTLEERLDKACEPGVDYVYKTRLVKQVLSNDFDE 1560
 Db 1501 KLNKLCRDELCRAEENCFIQKSDDKVTLEERLDKACEPGVDYVYKTRLVKQVLSNDFDE 1560
 QY 1561 YMAIEOTIKSGDEQVQGOQRTFISPIKCREALKLEKXKHYLMWGLSSDFWGEKPNLSY 1620
 Db 1561 YMAIEOTIKSGDEQVQGOQRTFISPIKCREALKLEKXKHYLMWGLSSDFWGEKPNLSY 1620
 QY 1621 IIGKDTWVHWPDEECQDBENOKQODLGAPTESVWVFGCPN 1663
 Db 1621 IIGKDTWVHWPDEECQDBENOKQODLGAPTESVWVFGSPN 1663

RESULT 11

ADB90023 standard; protein; 1663 AA.

AC ADB90023;

XX 04-DEC-2003 (first entry)

DT House complement component C3.

DE Human; antisense; complement component C3; inflammation; septic shock;

KW multiple organ failure; hyperacute organ failure; autoimmune disorder;

KW CNS inflammation; multiple sclerosis; atherosclerosis; tumour.

XX Homo sapiens.

OS US2003096775-A1.

PN 22-MAY-2003.

PD 23-OCT-2001; 2001US-00001076.

PF 23-OCT-2001; 2001US-00001076.

PR (ISIS-) ISIS PHARM INC.

XX Graham MJ, Watt AT;

PI N-PSDB; ADB9847.

XX MPI; 2003-606441/57.

DR N-PSDB; ADB9847.

XX New antisense oligonucleotides targeted to a nucleic acid molecule

PT encoding complement component C3, useful for treating a disease or

PT condition associated with complement component C3, e.g. autoimmune

PT disorder or infection.

XX Example 13; Page 28-34; 72pp; English.

PS The invention relates to a compound 8-50 nucleobases in length targeted

CC to a nucleic acid molecule encoding complement component C3. The compound

CC specifically hybridises with the nucleic acid molecule encoding

CC complement component C3 and inhibits the expression of complement

CC component C3, or specifically hybridises with at least an 8-nucleobase

CC portion of an active site on a nucleic acid molecule encoding complement

CC component C3. Also included are a composition comprising the compound and
 CC a pharmaceutical carrier or diluent, inhibiting the expression of
 CC complement component C3 in cells or tissues (comprising contacting the
 CC cells or tissues with the compound cited above) and treating an animal
 CC having a disease or condition associated with complement component C3
 CC comprising administering to the animal the compound cited above so that
 CC expression of complement component C3 is inhibited. The antisense
 CC compounds are useful for inhibiting the expression of complement
 CC component C3 in cells or tissues, or for treating an animal having a
 CC disease or condition associated with complement component C3 such as an
 CC autoimmune disorder (e.g. multiple sclerosis), an infection or
 CC atherosclerosis, inflammation, septic shock, multiple organ failure,
 CC hyperacute organ failure and CNS inflammation. The compounds are also
 CC useful as research reagents and diagnostics, in distinguishing functions
 CC of various members of a biological pathway, or for preventing or delaying
 CC infection, inflammation or tumour formation. The present sequence is a
 CC complement component C3 protein sequence.
 XX
 XX Sequence 1663 AA;

Query Match 22.98; Score 8598; DB 7; Length 1663;

Best Local Similarity 99.94; Pred. NO. 0;

Matches 1661; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGPTSGPSLLLLLTLLPLALGSPMYSIITPNILRESEETVLEAHDAQDQDPVTVTVH 60
 Db 1 MGPTSGPSLLLLLTLLPLALGSPMYSIITPNILRESEETVLEAHDAQDQDPVTVTVH 60
 QY 61 DPGKKLVLSSEKTVLPATNNGVNTFTIPANREPKSEKGNKVTVOATFGTVQVEKV 120
 Db 61 DPGKKLVLSSEKTVLPATNNGVNTFTIPANREPKSEKGNKVTVOATFGTVQVEKV 120
 QY 121 VLVSLSQGYLFIQDKTYITPGSTVLYRIFTVNNHKLPLVGRVTVMVNIENPESGIPVKQDSL 180
 Db 121 VLVSLSQGYLFIQDKTYITPGSTVLYRIFTVNNHKLPLVGRVTVMVNIENPESGIPVKQDSL 180
 QY 181 SSQNLGVLPMSWDIPELVNMGWKIRAYVENSPOOVSTEFVEKVEYLPFEVIVETE 240
 Db 181 SSQNLGVLPMSWDIPELVNMGWKIRAYVENSPOOVSTEFVEKVEYLPFEVIVETE 240
 QY 241 KFYIYINKEGLEVTITARFLYKKGVEGTAFVIFGIDQGEQRISLSPESLKRPIEDGSEV 300
 Db 241 KFYIYINKEGLEVTITARFLYKKGVEGTAFVIFGIDQGEQRISLSPESLKRPIEDGSEV 300
 QY 301 VLSRKVLDDGVQNPRAEDLVGKSLVYSATVILHSGSDMVQAESGIPITVSPYQIHFTKT 360
 Db 301 VLSRKVLDDGVQNPRAEDLVGKSLVYSATVILHSGSDMVQAESGIPITVSPYQIHFTKT 360
 QY 361 PKYFKPGMPFDLVFTVTPNPDGSPAYRVPVAVQGEDTVQSLTQGDGVAKLSINTHPSQKPL 420
 Db 361 PKYFKPGMPFDLVFTVTPNPDGSPAYRVPVAVQGEDTVQSLTQGDGVAKLSINTHPSQKPL 420
 QY 421 SITVTRTKKQELSEAEQATRTMQLPYSTVGNNSNYLHLSVLTETLRPGETLNVPFLRMD 480
 Db 421 SITVTRTKKQELSEAEQATRTMQLPYSTVGNNSNYLHLSVLTETLRPGETLNVPFLRMD 480
 QY 481 RAHEAKIRYTYLIMNKGRLIKAGROVRPQGDVLVPLSITTDIPISFRIVAYITLIGA 540
 Db 481 RAHEAKIRYTYLIMNKGRLIKAGROVRPQGDVLVPLSITTDIPISFRIVAYITLIGA 540
 QY 541 SGQREVVDVWVDVKDSCVGSILVKSQSDRQVPVQGMQLKIEGDHGRVVLVAVDK 600
 Db 541 SGQREVVDVWVDVKDSCVGSILVKSQSDRQVPVQGMQLKIEGDHGRVVLVAVDK 600
 QY 601 GVFLNKNKLTOSKIDWVVEKADICTPGSKGVYAGVSDAGLITFTSSSQQAQRAEL 660
 Db 601 GVFLNKNKLTOSKIDWVVEKADICTPGSKGVYAGVSDAGLITFTSSSQQAQRAEL 660
 QY 661 QCPQPAARRRRSVQLTEKMDKVGKYPKELRKCCEGDMRENPMRFSQCRRTFRFISLGEAC 720
 Db 661 QCPQPAARRRRSVQLTEKMDKVGKYPKELRKCCEGDMRENPMRFSQCRRTFRFISLGEAC 720
 QY 721 KKVPFLDCCNVTILRRQHAASHLGLARNLDEDDIIAENIVSRSEFFESLWNVDELKE 780

```

Db 721 KAVFLDCCNYITELRQARASHGLKASLNDDIABENIVSRSEFFPSWLNWVDDLKE 780
Qy 781 PRNGISTKLMNIFKDSITTWELAVMSDKGICVADPFVTVNQDFFIDLRLPYSV 840
Db 781 PRNGISTKLMNIFKDSITTWELAVMSDKGICVADPFVTVNQDFFIDLRLPYSV 840
Qy 841 RNEQVEIRAVLYNQRELKVRVELLHNPFAFSLATTKRHOOITIPKSSLSVPYVI 900
Db 841 RNEQVEIRAVLYNQRELKVRVELLHNPFAFSLATTKRHOOITIPKSSLSVPYVI 900
Qy 901 VPLKTGLQEVKAAVYHIFTSDGVRKSLKVPEGIRMNKTAVRTLDPERLGRGVQKE 960
Db 901 VPLKTGLQEVKAAVYHIFTSDGVRKSLKVPEGIRMNKTAVRTLDPERLGRGVQKE 960
Qy 961 DIPPADLSQDVPDTESETRILLQGTTPVAQMTEDAVDAERLKLIVTPSGCGEQNMIGTP 1020
Db 961 DIPPADLSQDVPDTESETRILLQGTTPVAQMTEDAVDAERLKLIVTPSGCGEQNMIGTP 1020
Qy 1021 TVIAVHYLDETEQNEKLEKROGALBLIKKGYTOOLAPROPSAFAAFVKAPSTWLT 1080
Db 1021 TVIAVHYLDETEQNEKLEKROGALBLIKKGYTOOLAPROPSAFAAFVKAPSTWLT 1080
Qy 1081 YVVKVFLAVNLIAIDSVLCGAVKWLILEKQKPDGVFOEDAPVHQMIGILRNNNEK 1140
Db 1081 YVVKVFLAVNLIAIDSVLCGAVKWLILEKQKPDGVFOEDAPVHQMIGILRNNNEK 1140
Qy 1141 MALTAFLVLSIQEAKDICEQVNSLPGSITKAGDFLEANYMTLORSYTVIAIAGYALQNG 1200
Db 1141 MALTAFLVLSIQEAKDICEQVNSLPGSITKAGDFLEANYMTLORSYTVIAIAGYALQNG 1200
Qy 1201 RLKGPLLNKFTTAKDNKRWEDPGKQLYNVEATSYALLALLQKDFPVPVVRMLNEQR 1260
Db 1201 RLKGPLLNKFTTAKDNKRWEDPGKQLYNVEATSYALLALLQKDFPVPVVRMLNEQR 1260
Qy 1261 YVGGYGTQATFWFQALAYQKADPHQELNDVSLQPLPSRSSKITTHRIHWSASLLR 1320
Db 1261 YVGGYGTQATFWFQALAYQKADPHQELNDVSLQPLPSRSSKITTHRIHWSASLLR 1320
Qy 1321 SEETKENGFTVTAEGKGQGLSVVTVYHAKADQLTCKNFKDLKVIKPAETEKRPQDA 1380
Db 1321 SEETKENGFTVTAEGKGQGLSVVTVYHAKADQLTCKNFKDLKVIKPAETEKRPQDA 1380
Qy 1381 KNTMLEICTRYRGDQATMSILDI SMTGFPDPTDLKOLANGVDVRYISKVELDKAFSD 1440
Db 1381 KNTMLEICTRYRGDQATMSILDI SMTGFPDPTDLKOLANGVDVRYISKVELDKAFSD 1440
Qy 1441 RNTLIYLDKVSHEDDCLAFKHQYFNVELIQGAVKVYAYYNLSEESTRFVHPKSDG 1500
Db 1441 RNTLIYLDKVSHEDDCLAFKHQYFNVELIQGAVKVYAYYNLSEESTRFVHPKSDG 1500
Qy 1501 KLNKLCRDELRCRAEENCFIQSDDKVTLERLDKACEPGVDVYVYKTRLVKQLSNDPDE 1560
Db 1501 KLNKLCRDELRCRAEENCFIQSDDKVTLERLDKACEPGVDVYVYKTRLVKQLSNDPDE 1560
Qy 1561 YMAIEQTIKSGSDEVQVQOQRTFISPKREALKULEKHVYLMWGLSDFGWGKPNLSY 1620
Db 1561 YMAIEQTIKSGSDEVQVQOQRTFISPKREALKULEKHVYLMWGLSDFGWGKPNLSY 1620
Qy 1621 IIGKDTWVHWPEDDECODENOKQCDLGAFTESMVVFGCPN 1663
Db 1621 IIGKDTWVHWPEDDECODENOKQCDLGAFTESMVVFGCPN 1663

```

RESULT 12

ID ADD93518 standard; protein; 1663 AA.

XX AC

XX ADD93518;

XX 29-JAN-2004 (first entry)

XX Novel NOV1a, homologous to human complement C3 precursor.

DE

XX NOV1a; human; complement C3; gene therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Peptide 1..22

XX Protein /note="Signal peptide"

XX /note="Maure protein"

XX W02003078572-A2.

XX 25-SEP-2003.

XX 06-MAR-2003; 2003WO-US006859.

XX 15-MAR-2002; 2002US-0365034P.

XX 19-MAR-2002; 2002US-0365477P.

XX 21-MAR-2002; 2002US-0366420P.

XX 05-MAR-2003; 2003US-00379747.

XX (CURA-) CURAGEN CORP.

XX Burgess CE, Chant JS, Chaudhuri A, Edinger SR, Gangolli EA;

XX Malyankar UM, Miller CE, Ooi CE, Ort T, Patturajan M, Rastelli L;

XX Rieger DK, Shinkets RA, Zerhusen BD;

XX WPI; 2003-779122/73.

XX N-PSDB; ADD93517.

XX New isolated NOVX polypeptides and polynucleotides, useful for

XX preventing, diagnosing or treating NOVX-associated disorders, e.g.

XX osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,

XX asthma, or infections.

XX Claim 1; Page 100; 205pp; English.

XX The present sequence is the protein sequence of a novel polypeptide,

XX designated NOV1a, that shows 99% sequence homology to the human

XX complement C3 precursor. The invention is based on the identification of

XX proteins and polypeptides, and the nucleic acids encoding them, that are

XX differentially modulated in a pathological state, disease or an abnormal

XX condition or state. These are targets for therapeutic agents and can be

XX used in screening methodologies to identify candidate therapeutic agents

XX which interact with the target and thereby exert a desired or favourable

XX effect, e.g. in neurogenesis, cell differentiation, cell proliferation,

XX haematopoiesis, wound healing and angiogenesis. Methods for diagnosis,

XX treatment and prevention of disorders involving the novel human nucleic

XX acids and proteins are provided. The polypeptides can also be used to

XX raise antibodies useful e.g. in diagnosis and therapy.

XX Sequence 1663 AA;

Query Match 99.9%; Score 8598; DB 7; Length 1663;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1661; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGPTSGPSLLLLTHLPALGSPMYSITPNILRSEETWVLEAHDAGQDVPVTVVH 60

Db 1 MGPTSGPSLLLLTHLPALGSPMYSITPNILRSEETWVLEAHDAGQDVPVTVVH 60

Qy 61 DPEGKLVLSSEKTVLTATNMGNVTITIPANREFKSEKGRNKFVTVQATFGTVVVKV 120

Db 61 DPEGKLVLSSEKTVLTATNMGNVTITIPANREFKSEKGRNKFVTVQATFGTVVVKV 120

Qy 121 VLVSQSGVLFITQDKTIYTPGSTVLYRIFTVNKKLLPVGRTVMVNIENPEGIPVKQDSL 180

Db 121 VLVSQSGVLFITQDKTIYTPGSTVLYRIFTVNKKLLPVGRTVMVNIENPEGIPVKQDSL 180

Qy 181 SSONQLGVLPISWDIPELVNMQWIRAYENSPOQVFSTFEVKEYVLPSPFVIVPTE 240

Db 181 SSONQLGVLPISWDIPELVNMQWIRAYENSPOQVFSTFEVKEYVLPSPFVIVPTE 240

QY 241 KFYIYNEKLEVTITARFLYKKGVEGTAFVIFGIQDGEORISLPSLKRPIEDSGSEV 300
DB 241 KFYIYNEKLEVTITARFLYKKGVEGTAFVIFGIQDGEORISLPSLKRPIEDSGSEV 300
QY 301 VLSKVLDDGVQNPRAEDLVGKSLYVSATVILHSGSDMVQABRSQIPVTSFYQIHFTKT 360
DB 301 VLSKVLDDGVQNPRAEDLVGKSLYVSATVILHSGSDMVQABRSQIPVTSFYQIHFTKT 360
QY 361 PKYKPKMPDLVGVFVTPNPDGSPAYRVPVAVOGEOTVQSLTQGDGVAKLSTNTHPSOKPL 420
DB 361 PKYKPKMPDLVGVFVTPNPDGSPAYRVPVAVOGEOTVQSLTQGDGVAKLSTNTHPSOKPL 420
QY 421 SITVTKKQELSEAEQATRMQALPYSTVGNNSNYLHLSVLRTLPGETLNVLNPLRMD 480
DB 421 SITVTKKQELSEAEQATRMQALPYSTVGNNSNYLHLSVLRTLPGETLNVLNPLRMD 480
QY 481 RAHEAKRYTYTILMKNKGRLLKAGROVREPQDGLVPLSITTDPIPSFRILVAYYTLIGA 540
DB 481 RAHEAKRYTYTILMKNKGRLLKAGROVREPQDGLVPLSITTDPIPSFRILVAYYTLIGA 540
QY 541 SGQREVVADSVWVDVWDSVGSILVKSQSEDRQVPVQGMVTLKIEGDHGRVAVLVADK 600
DB 541 SGQREVVADSVWVDVWDSVGSILVKSQSEDRQVPVQGMVTLKIEGDHGRVAVLVADK 600
QY 601 GVFLVKNKLTOSKIWDVVEKADICTPGSGKDVAGVFSAGLTFTSSSGOQTAQRAEL 660
DB 601 GVFLVKNKLTOSKIWDVVEKADICTPGSGKDVAGVFSAGLTFTSSSGOQTAQRAEL 660
QY 661 QCPQPAARRRSVQLTEKMDKVKYKPKELRKCCEGRENPMRPSQORRTFRFSLGEAC 720
DB 661 QCPQPAARRRSVQLTEKMDKVKYKPKELRKCCEGRENPMRPSQORRTFRFSLGEAC 720
QY 721 KKVFLDCNVTILRQHARASHLGLARSLNLDIIAENIVSRSEFFESMLWNVDELKE 780
DB 721 KKVFLDCNVTILRQHARASHLGLARSLNLDIIAENIVSRSEFFESMLWNVDELKE 780
QY 781 PPKNGISTKLMNIFLKDSITTWELAVSMDKKGICVADPEVTVNQDFFIDLPLPSVY 840
DB 781 PPKNGISTKLMNIFLKDSITTWELAVSMDKKGICVADPEVTVNQDFFIDLPLPSVY 840
QY 841 RNEQVEIRAVLYNQKQELKRVVELLHNPAPCSLATTKRHQQTITIPPKSSISVPIVI 900
DB 841 RNEQVEIRAVLYNQKQELKRVVELLHNPAPCSLATTKRHQQTITIPPKSSISVPIVI 900
QY 901 VPLKTGQVEVKAAYVHHFISDGVKSLKVVPGEIRMNKTVAVPTLDPERLREGVQKE 960
DB 901 VPLKTGQVEVKAAYVHHFISDGVKSLKVVPGEIRMNKTVAVPTLDPERLREGVQKE 960
QY 961 DIPPADLSQVDPTESETRILLQGTVEAQMTEDAVDAERLKLIVTPSGCGEQNNIGMTP 1020
DB 961 DIPPADLSQVDPTESETRILLQGTVEAQMTEDAVDAERLKLIVTPSGCGEQNNIGMTP 1020
QY 1021 TVIAVHVLDETEQWKEGKROGALIKKGTQQLAPROPSAFAAFVRASTWLT 1080
DB 1021 TVIAVHVLDETEQWKEGKROGALIKKGTQQLAPROPSAFAAFVRASTWLT 1080
QY 1081 YVVKVFLAVNLIAIDSVQLCGAVKWLILEKQKPDGVFOEDAPVTHQEMIGLNNNEKD 1140
DB 1081 YVVKVFLAVNLIAIDSVQLCGAVKWLILEKQKPDGVFOEDAPVTHQEMIGLNNNEKD 1140
QY 1141 MALTAFLVLSIQAKDICEBOVNSLPGSITTKAGDPLEANVMYLNORSYTVAGVLAQMG 1200
DB 1141 MALTAFLVLSIQAKDICEBOVNSLPGSITTKAGDPLEANVMYLNORSYTVAGVLAQMG 1200
QY 1201 RLKGFLLNKFLTTAKQKRWEDSGKQLYNVEATSTALLALQLKDPDFVPPVFWLNEQR 1260
DB 1201 RLKGFLLNKFLTTAKQKRWEDSGKQLYNVEATSTALLALQLKDPDFVPPVFWLNEQR 1260
QY 1261 YVGGYGSTQATFWFOALAQYQKADPDHQLNLDVSLQLPSRSSKITHRIHWSASLLR 1320
DB 1261 YVGGYGSTQATFWFOALAQYQKADPDHQLNLDVSLQLPSRSSKITHRIHWSASLLR 1320

QY 1321 SEETKENEGETVTAEGKGQGLSVWTMYHAKAKQDLTKNFKDLKVTIKPAPETEKRPQDA 1380
DB 1321 SEETKENEGETVTAEGKGQGLSVWTMYHAKAKQDLTKNFKDLKVTIKPAPETEKRPQDA 1380
QY 1381 KNTWILICTRYRGDQDATMSILDISMTGFAPDPTDQLKQLANGVDRYISKYELDKAFSD 1440
DB 1381 KNTWILICTRYRGDQDATMSILDISMTGFAPDPTDQLKQLANGVDRYISKYELDKAFSD 1440
QY 1441 RNTLIIYLDKVSSEDDCLAFKVOYFNVELIQPQAVKYVAYNLEBSCTRFYHPEKEDG 1500
DB 1441 RNTLIIYLDKVSSEDDCLAFKVOYFNVELIQPQAVKYVAYNLEBSCTRFYHPEKEDG 1500
QY 1501 KLNLCRDELRCRAENCFIQKSDDKVTLLEELDKACBEGVDYVYKTLVKVQLSNDPDE 1560
DB 1501 KLNLCRDELRCRAENCFIQKSDDKVTLLEELDKACBEGVDYVYKTLVKVQLSNDPDE 1560
QY 1561 YIMAEQITIKSGSDEVQVQGOQRTFISPIKCREALKLEKKHYLMWGLSSDFWGEKPNLSY 1620
DB 1561 YIMAEQITIKSGSDEVQVQGOQRTFISPIKCREALKLEKKHYLMWGLSSDFWGEKPNLSY 1620
QY 1621 IIGKDTVWEHWPEDDECOEENKQCOQDILGAFTESMVVFGCPN 1663
DB 1621 IIGKDTVWEHWPEDDECOEENKQCOQDILGAFTESMVVFGCPN 1663

RESULT 13

ADK12322
ID ADK12322 standard; protein; 1663 AA.

XX AC ADK12322;

XX XX 20-MAY-2004 (first entry)

XX XX Human complement component C3.

XX KW Antisense therapy; human; complement component C3; autoimmune disorder;
multiple sclerosis; infection; atherosclerosis; neuroprotective;
XX KW antiatherosclerotic; antimicrobial; antiinflammatory; cytostatic.

XX OS Homo sapiens.

XX XX US2004043956-A1.

XX XX 04-MAR-2004.

XX XX 18-AUG-2003; 2003US-00642802.

XX XX 23-OCT-2001; 2001US-00001076.

XX XX (GRAH/) GRAHAM M J.

XX XX (WATT/) WATT A T.

XX XX Graham MJ, Watt AT;

XX XX WPI: 2004-225730/21.

XX XX N-PSDS; ADK12146.

XX XX New antisense compound targeted to a nucleic acid molecule encoding
complement component C3, useful for treating multiple sclerosis, an
infection or atherosclerosis.

XX XX Disclosure; Page 31-36; 74pp; English.

XX XX The present invention relates to antisense compounds targeted to a
nucleic acids encoding human and mouse complement component C3. The
antisense compound comprises an antisense oligonucleotide that
specifically hybridizes with the nucleic acid and inhibits the expression
of complement component C3 in cells. The antisense oligonucleotide is a
chimeric oligonucleotide. The antisense oligonucleotide comprises at
least one modified internucleoside linkage, preferably a phosphorothioate
linkage. It also comprises at least one modified sugar moiety, preferably
a 2'-O-methoxyethyl (2'-MOE) sugar moiety. The antisense oligonucleotide
further comprises at least one modified nucleobase, preferably a 5-

CC methionylcysteine. The antisense oligonucleotides are useful for the
 CC treatment of diseases such as autoimmune disorders e.g. multiple
 CC sclerosis, infections, and atherosclerosis. The present sequence
 CC represents human complement component C3.

XX
 SQ Sequence 1663 AA;

Query Match 99.9%; Score 8598; DB 8; Length 1663;
 Best Local Similarity 99.9%; Pred. No 0;
 Matches 1663; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGPTSGSLLLLLLTHPLALGSPMYSIITPNILRESEETVWLEAHDAQSDVPVTVH 60
 Db 1 MGPTSGSLLLLLLTHPLALGSPMYSIITPNILRESEETVWLEAHDAQSDVPVTVH 60

Qy 61 DFPKGLVLSSEKTLVTPATNMGNVTFTIPANREFKSEKGNKFTVQATFGTOVBEK 120
 Db 61 DFPKGLVLSSEKTLVTPATNMGNVTFTIPANREFKSEKGNKFTVQATFGTOVBEK 120

Qy 121 VLVSQGYLFIQTDKTIYTPGSTVLYRFTVNHKLLPVGRVVMNINPESGIPVKQSL 180
 Db 121 VLVSQGYLFIQTDKTIYTPGSTVLYRFTVNHKLLPVGRVVMNINPESGIPVKQSL 180

Qy 181 SSQNLGVLPDWDIPELVNMGWKIRAYENSPOQVSTEFVEKEYVLPSEVIIVPE 240
 Db 181 SSQNLGVLPDWDIPELVNMGWKIRAYENSPOQVSTEFVEKEYVLPSEVIIVPE 240

Qy 241 KFYIYNEKLEVTITARFLYKKGVEGTAFVIFGIQDGEORSLPESLKRPIEDSGEV 300
 Db 241 KFYIYNEKLEVTITARFLYKKGVEGTAFVIFGIQDGEORSLPESLKRPIEDSGEV 300

Qy 301 VLSKVLVDGQVPRADLVGKSLYSATVILHSGSDMVOAERSGIPVTSPYQHFTK 360
 Db 301 VLSKVLVDGQVPRADLVGKSLYSATVILHSGSDMVOAERSGIPVTSPYQHFTK 360

Qy 361 PKYFKQMPDLVFTNPDGSPAYRVPVAVQGEDTVQSLTQDGVAKLSINTHPSQPL 420
 Db 361 PKYFKQMPDLVFTNPDGSPAYRVPVAVQGEDTVQSLTQDGVAKLSINTHPSQPL 420

Qy 421 SITVTKKQSLSEABQATRMQALPSTVGNSSNNYLHLSVLTARPGETLNVNPLRMD 480
 Db 421 SITVTKKQSLSEABQATRMQALPSTVGNSSNNYLHLSVLTARPGETLNVNPLRMD 480

Qy 481 RAHEAKIRYTYLIMNKGRLKAGROVRPQDQVPLVLSITTDIPFRLVAYTYLIGA 540
 Db 481 RAHEAKIRYTYLIMNKGRLKAGROVRPQDQVPLVLSITTDIPFRLVAYTYLIGA 540

Qy 541 SGQREVADSVWVDVKDSCVGLVWKSQSEDRQPVPGQMTLKEGDHGARVILVADK 600
 Db 541 SGQREVADSVWVDVKDSCVGLVWKSQSEDRQPVPGQMTLKEGDHGARVILVADK 600

Qy 601 GFVLNKKNLTKSILWVVEKADIGCTPGSGKHGYAGVPSDAGLFTSSSGQQTARAE 660
 Db 601 GFVLNKKNLTKSILWVVEKADIGCTPGSGKHGYAGVPSDAGLFTSSSGQQTARAE 660

Qy 661 QCPOPARRRRSVQLTEKMDKVGKYPKELRKCCEDGMENPMRPSQCORRTFISLGAC 720
 Db 661 QCPOPARRRRSVQLTEKMDKVGKYPKELRKCCEDGMENPMRPSQCORRTFISLGAC 720

Qy 721 KKVPLDCCNITELRQPARASHLGLARNLDEDIABENIVSRSEFPESMLWNVEDLKE 780
 Db 721 KKVPLDCCNITELRQPARASHLGLARNLDEDIABENIVSRSEFPESMLWNVEDLKE 780

Qy 781 PPKNIGISTKLMNIFLKDSITTWELAVSMDSKXGICVADPFVTVMQPFFIDRLRPSVW 840
 Db 781 PPKNIGISTKLMNIFLKDSITTWELAVSMDSKXGICVADPFVTVMQPFFIDRLRPSVW 840

Qy 841 RNEQVEIRAVLYNQROELKVRVELLHNPAPCSLATTGRHQCTITTPPKSSLSVPYVI 900
 Db 841 RNEQVEIRAVLYNQROELKVRVELLHNPAPCSLATTGRHQCTITTPPKSSLSVPYVI 900

Qy 901 VPLKGTQGEVEKAAVYHHFISDGVKSLKVPPEGIRNMKNKTVAVRTLDPERLREGVOKE 960
 Db 901 VPLKGTQGEVEKAAVYHHFISDGVKSLKVPPEGIRNMKNKTVAVRTLDPERLREGVOKE 960

Db 901 VPLKGTQGEVEKAAVYHHFISDGVKSLKVPPEGIRNMKNKTVAVRTLDPERLREGVOKE 960
 Qy 961 DIPPADLSQVDPDTESETRILLQGPVAMTDEDAERLKLIVTSPGCGEQNMIGWTP 1020
 Db 961 DIPPADLSQVDPDTESETRILLQGPVAMTDEDAERLKLIVTSPGCGEQNMIGWTP 1020

Qy 1021 TVIAVHYLDETEQWEKFGLEKQKQGALELIKKGYTQQLAFROPSSAFAAFVKRAPSTWLT 1080
 Db 1021 TVIAVHYLDETEQWEKFGLEKQKQGALELIKKGYTQQLAFROPSSAFAAFVKRAPSTWLT 1080

Qy 1081 YVYVPSFLAVNLIADSOVLGAVKWLILEKQKQDGVFOEDAPVHOBMIQGLNNNEKD 1140
 Db 1081 YVYVPSFLAVNLIADSOVLGAVKWLILEKQKQDGVFOEDAPVHOBMIQGLNNNEKD 1140

Qy 1141 MALTAFLVLSLOEAKDICEEQVNSLPGSITKAGDFLEANYMNLORSYTVIAGYALAQMG 1200
 Db 1141 MALTAFLVLSLOEAKDICEEQVNSLPGSITKAGDFLEANYMNLORSYTVIAGYALAQMG 1200

Qy 1201 RLKGPLLNKFLITAKRWEDPGKOLYNVEATSYALLALLQLKDFDPVPPVWLNNEOR 1260
 Db 1201 RLKGPLLNKFLITAKRWEDPGKOLYNVEATSYALLALLQLKDFDPVPPVWLNNEOR 1260

Qy 1261 YVGGYVGSQTQATFMVVFQALAQYQKADAPDHQELNLDVSLQLPSSRSKITHRIHWEASLLR 1320
 Db 1261 YVGGYVGSQTQATFMVVFQALAQYQKADAPDHQELNLDVSLQLPSSRSKITHRIHWEASLLR 1320

Qy 1321 SEETKENEGFTVTAEGKQGTLSVVTMYHAKAQDLTCNKFDLVTKIPAPETEKRPQDA 1380
 Db 1321 SEETKENEGFTVTAEGKQGTLSVVTMYHAKAQDLTCNKFDLVTKIPAPETEKRPQDA 1380

Qy 1381 KNTMILEICTRYRGQDQATMSILDISMGTGAPDQDLKQLANGVDYRYSKYELDKAFSD 1440
 Db 1381 KNTMILEICTRYRGQDQATMSILDISMGTGAPDQDLKQLANGVDYRYSKYELDKAFSD 1440

Qy 1441 RNTLIYLDKVSHEDDCLAPKHQYFNVELIQGAVKAYVYNNLESCCTFYHPEKEDG 1500
 Db 1441 RNTLIYLDKVSHEDDCLAPKHQYFNVELIQGAVKAYVYNNLESCCTFYHPEKEDG 1500

Qy 1501 KLKLCRDELRCABEENCPIQKSDDKVLEERLDKACEPGVDYVYKTRLVKQVLSNDFDE 1560
 Db 1501 KLKLCRDELRCABEENCPIQKSDDKVLEERLDKACEPGVDYVYKTRLVKQVLSNDFDE 1560

Qy 1561 YNMAIECTIKSDEVOVQOQRTFISPIKCPREALKLEKKHYLMKGLSSDQWGEKPNLSY 1620
 Db 1561 YNMAIECTIKSDEVOVQOQRTFISPIKCPREALKLEKKHYLMKGLSSDQWGEKPNLSY 1620

Qy 1621 IIGKDTWVEHWPEDDEECQDEENKQCDLGAFTESMVVFQCPN 1663
 Db 1621 IIGKDTWVEHWPEDDEECQDEENKQCDLGAFTESMVVFQCPN 1663

RESULT 14

ADN04780
 ID ADN04780 standard; protein; 1663 AA.

XX AC ADN04780;

XX DT 01-JUL-2004 (first entry)

XX DE Antisporiatric protein sequence #570.

XX KW antipsoriatric; gene therapy; psoriasis; diagnosis.

XX OS Homo sapiens.

XX PN WO2004028479-A2.

XX PD 08-APR-2004.

XX PF 25-SEP-2003; 2003WC-US030907.

XX PR 25-SEP-2002; 2002US-0414006P.

XX

PA (GETH) GENENTECH INC.
 XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WT;
 PI Wu TD;
 XX WPI; 2004-305105/28.
 DR N-PSDB; ADN04779.
 XX New PRO nucleic acid or polypeptide, useful for preparing a
 PT pharmaceutical composition for diagnosing or treating psoriasis in a
 PT mammal.
 XX Claim 9; SEQ ID NO 1174; 3069pp; English.
 XX The invention relates to novel polynucleotide and polypeptides for
 CC treating psoriasis or a sequence having at least 80% identity to the
 CC above sequences. The nucleic acid is useful for preparing a composition
 CC for diagnosing or treating psoriasis in a mammal. This sequence
 CC corresponds to one of the polypeptides of the invention.
 XX Sequence 1663 AA;
 SQ

Query Match 99.9%; Score 8598; DB 8; Length 1663;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1661; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGPTSGPSLLLLTHLPLALGSPMYSIITPNILRESEETMVLHAHQDQVPTVTYH 60
 Db 1 MGPTSGPSLLLLTHLPLALGSPMYSIITPNILRESEETMVLHAHQDQVPTVTYH 60

Qy 61 DPEGKLVLSSEKTVLTPATNMGVNTFTIPANREFKSEKGRNKFVTVQATFGQVVEKV 120
 Db 61 DPEGKLVLSSEKTVLTPATNMGVNTFTIPANREFKSEKGRNKFVTVQATFGQVVEKV 120

Qy 121 VLVSQSGYLFIQTDKTIYTPGSTVLYRIFTVNHKLLFVGRVTVMVNIENPEGIPVKQDSL 180
 Db 121 VLVSQSGYLFIQTDKTIYTPGSTVLYRIFTVNHKLLFVGRVTVMVNIENPEGIPVKQDSL 180

Qy 181 SSQNLGVLPNDIPELNMGQWIRAYYENSPOOVSTEFVEKVEYVLPSEFVIEPTE 240
 Db 181 SSQNLGVLPNDIPELNMGQWIRAYYENSPOOVSTEFVEKVEYVLPSEFVIEPTE 240

Qy 241 KFYIYNKGLVITITARELYGKVEGTAFVIFGIDGQEQRIISLPESLKRPIEDSGSEV 300
 Db 241 KFYIYNKGLVITITARELYGKVEGTAFVIFGIDGQEQRIISLPESLKRPIEDSGSEV 300

Qy 301 VLSRKVLGVQNPRAEDLVGSLVSAIVLHSGSDMVQAESGIPVTSPIQIHFTKT 360
 Db 301 VLSRKVLGVQNPRAEDLVGSLVSAIVLHSGSDMVQAESGIPVTSPIQIHFTKT 360

Qy 361 PKYFKGMPFDMVFTNPDGSPAYRVPVAVQGEDTVQSLTQGGGVAKLSINTHPSQKPL 420
 Db 361 PKYFKGMPFDMVFTNPDGSPAYRVPVAVQGEDTVQSLTQGGGVAKLSINTHPSQKPL 420

Qy 421 SIIVRTKQELSEABQATMQLPYSIVTGNNSNVLHLSVLTRELPGETLNNFLRMD 480
 Db 421 SIIVRTKQELSEABQATMQLPYSIVTGNNSNVLHLSVLTRELPGETLNNFLRMD 480

Qy 481 RAHEAKIRYTYLIMNKGRLKAGQVREPQGLVPLSITTDPIPSFLVAYYTLIGA 540
 Db 481 RAHEAKIRYTYLIMNKGRLKAGQVREPQGLVPLSITTDPIPSFLVAYYTLIGA 540

Qy 541 SGQREVVDVVDVVDVVDVVDVVDVVDVVDVVDVVDVVDVVDVVDVVDVVDVVDV 600
 Db 541 SGQREVVDVVDVVDVVDVVDVVDVVDVVDVVDVVDVVDVVDVVDVVDVVDVVDV 600

Qy 601 GVFLVNNKNTQSKLWDVVEKADTCTPGSGKVAGVFSADGLTFTSSSQQQAQAE 660
 Db 601 GVFLVNNKNTQSKLWDVVEKADTCTPGSGKVAGVFSADGLTFTSSSQQQAQAE 660

Qy 661 QCQPAARRRRSVQLTEKMDVGVKYPKELKCCEDGMRENPMRFSQRRTRFISLGEAC 720
 Db 661 QCQPAARRRRSVQLTEKMDVGVKYPKELKCCEDGMRENPMRFSQRRTRFISLGEAC 720

Qy 721 KKVFLDCNCVITELRRQHARASHLGLARNLNLEDIIAENIVSRSEFPESIMWNVEDLKE 780
 Db 721 KKVFLDCNCVITELRRQHARASHLGLARNLNLEDIIAENIVSRSEFPESIMWNVEDLKE 780

Qy 781 PPKNGISTKLMNIFLKDSITTWELAVSMDSKKGICVADPFEVTVMQDFFIDLRPLYSV 840
 Db 781 PPKNGISTKLMNIFLKDSITTWELAVSMDSKKGICVADPFEVTVMQDFFIDLRPLYSV 840

Qy 841 RNEQVEIRAVLYNRYRQNELKVRVELLHNPAPCSLATTKRHQQTITIPPKSSLSVPYV 900
 Db 841 RNEQVEIRAVLYNRYRQNELKVRVELLHNPAPCSLATTKRHQQTITIPPKSSLSVPYV 900

Qy 901 VPLTGTQVEVKAAYVHHFISDGVKSLKAVPEGRMNKTVAVRTLPERLGRGVOKE 960
 Db 901 VPLTGTQVEVKAAYVHHFISDGVKSLKAVPEGRMNKTVAVRTLPERLGRGVOKE 960

Qy 961 DIPADLSQVPTDSESTRILLQGTTPVAQMTDANDAEKHLIVTPSCGGEQNNMIGTP 1020
 Db 961 DIPADLSQVPTDSESTRILLQGTTPVAQMTDANDAEKHLIVTPSCGGEQNNMIGTP 1020

Qy 1021 TVIAVHYLDTEQWEKRGLEKROGALIELIKGYTQOLAFRQPSAFAAFVKEAPSTWLTA 1080
 Db 1021 TVIAVHYLDTEQWEKRGLEKROGALIELIKGYTQOLAFRQPSAFAAFVKEAPSTWLTA 1080

Qy 1081 YVVKVFLAVNLIAIDSVQLCGAVKWLILEKQKPDGVFOEDAPVTHQEMIGGLRNNKED 1140
 Db 1081 YVVKVFLAVNLIAIDSVQLCGAVKWLILEKQKPDGVFOEDAPVTHQEMIGGLRNNKED 1140

Qy 1141 MALTAFLVLSIQAKDICEPOVNSLPGSITKAGDFLEANYMNLQRSYTVAIAGYALAQWG 1200
 Db 1141 MALTAFLVLSIQAKDICEPOVNSLPGSITKAGDFLEANYMNLQRSYTVAIAGYALAQWG 1200

Qy 1201 RLKGLPLNKFTTAKDKNRWEDPGKQIYNVEATSYALLALLQKDFDFVPPVVRWLNQOR 1260
 Db 1201 RLKGLPLNKFTTAKDKNRWEDPGKQIYNVEATSYALLALLQKDFDFVPPVVRWLNQOR 1260

Qy 1261 YVGGYGSTQATFWFQALAQYQKADPADHQLNLDVSLQPSRSSKI THRIHWESASLLR 1320
 Db 1261 YVGGYGSTQATFWFQALAQYQKADPADHQLNLDVSLQPSRSSKI THRIHWESASLLR 1320

Qy 1321 SEETKENEFTVTAEGKGQGTLSVVTWYHAKAKDQUTCNKFDLKVITIKPAPETEKRPDA 1380
 Db 1321 SEETKENEFTVTAEGKGQGTLSVVTWYHAKAKDQUTCNKFDLKVITIKPAPETEKRPDA 1380

Qy 1381 KNTMILICTRYRGDQDATMSILLDISMTGFAPDITDQLQLANGVDVRYISKYELDKAFSD 1440
 Db 1381 KNTMILICTRYRGDQDATMSILLDISMTGFAPDITDQLQLANGVDVRYISKYELDKAFSD 1440

Qy 1441 RNTLIILYLDKVSSEDDCLAFKHQVFNVELTOPGAKVYAYYNLEESCTRYHPEKEDG 1500
 Db 1441 RNTLIILYLDKVSSEDDCLAFKHQVFNVELTOPGAKVYAYYNLEESCTRYHPEKEDG 1500

Qy 1501 KLNKLCRDELRCRAEENCFIQKSDDKVLEERLDKACEPGVDYVYKTRLVKQLSNDDE 1560
 Db 1501 KLNKLCRDELRCRAEENCFIQKSDDKVLEERLDKACEPGVDYVYKTRLVKQLSNDDE 1560

Qy 1561 YMAIEQTIKSGSEDEVQVGGQRTFISPIKCREALKLEBKHYLMWGLSSDFWGEKPNLSY 1620
 Db 1561 YMAIEQTIKSGSEDEVQVGGQRTFISPIKCREALKLEBKHYLMWGLSSDFWGEKPNLSY 1620

Qy 1621 IIGKDTVWEHWPEDCEQDEENKQCDLGAFTESMVVFGCFN 1663
 Db 1621 IIGKDTVWEHWPEDCEQDEENKQCDLGAFTESMVVFGCFN 1663

RESULT 15
 AAW40988
 ID AAW40988 standard; protein; 1663 AA.
 XX
 AC AAW40988;
 XX
 DT 09-APR-1998 (first entry)

XX DE Human C3 protein mutant R1303X, R1320X.
 XX KW Human; C3 protein; convertase; complement pathway protein; infection;
 XX KW down-regulation resistant C3 convertase; xenograft rejection; therapy;
 XX KW complement-mediated disease; autoimmune disease; leukaemia cell; tumour;
 XX KW complement-mediated response; MHC-mismatched lymphocyte; mutain.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX FT Misc-difference 1303 /label= Glu, Tyr, Cys, Trp, Gln, Gly
 XX FT Misc-difference 1320 /label= Glu, Tyr, Cys, Trp, Gln, Gly
 XX FT
 XX PN WO9732981-A1.
 XX XX 12-SEP-1997.
 XX XX 04-MAR-1997; 97WO-GB000603.
 XX XX 07-MAR-1996; 96GB-00004865.
 XX XX 07-JUN-1996; 96GB-00011896.
 XX XX 08-JUL-1996; 96GB-00014293.
 XX XX 19-NOV-1996; 96GB-00024028.
 XX XX (IMUT-) IMUTRAN LTD.
 XX PA Farries TC, Harrison RA;
 XX PI WPI; 1997-457534/42.
 XX DR
 XX PT Modified complement pathway protein that forms C3 convertase resistant to
 XX PT down-regulation - used to exhaust the complement pathway by super-
 XX PT activation, especially for preventing graft rejection, etc.
 XX XX
 XX PS Claim 6; Page; 1232p; English.
 XX CC This sequence represents a mutated human C3 protein of the invention (see
 CC AA034606 for wild type protein). This protein is a protein of the
 CC invention, and is a modified native complement pathway protein (A) that
 CC forms a down-regulation resistant C3 convertase. (A), their variants,
 CC fragments and conjugates are used to deplete levels of complement pathway
 CC proteins (by superactivation until one or more components are exhausted),
 CC specifically to prevent rejection of foreign material (particularly a
 CC xenograft), but also to prevent complement-mediated diseases resulting
 CC from (surgical) injury or antibody-antigen interaction in autoimmune
 CC disease, also to localise and/or amplify endogenous complement protein
 CC conversion and deposition at a specific site (e.g. a virus, infected cell
 CC or tumour, to increase sensitivity to complement-mediated responses; a
 CC particular application is eliminating any cancer cells left after
 CC surgical removal of a tumour). Also contemplated is ex vivo treatment,
 CC especially by passing blood through a matrix containing (A) (this may
 CC remove additional anaphylactic peptides and other inflammatory mediators)
 CC or killing of leukaemia cells or MHC-mismatched lymphocytes in extracted
 CC bone marrow. Since (A) is not inhibited by factor I, it can bind
 CC repeatedly to factor B (which is then inactivated), causing inactivation
 CC of the alternative pathway by consumption of factor B
 XX SQ Sequence 1663 AA;
 Query Match 99.98; Score 8597; DB 2; Length 1663;
 Best Local Similarity 99.94; Pred. No. 0;
 Matches 1661; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 MGPTSGPSLLLLTLPLALGSPMYSIITPNILRLESEETMVLBAHQDQGVPTVTVH 60
 Db 1 MGPTSGPSLLLLTLPLALGSPMYSIITPNILRLESEETMVLBAHQDQGVPTVTVH 60
 Qy 61 DFPCKLVLSSEKTVLTPATNMGNTVTFIPANREFKSEKGNKFTVQATFGTOVEKV 120
 Db 61 DFPCKLVLSSEKTVLTPATNMGNTVTFIPANREFKSEKGNKFTVQATFGTOVEKV 120

Qy 121 VLVSQSGYLFIOTDKTIYTPGSTVLRYRIFTVNHKLLPVGRVTMWNINENPEGIPVKQDSL 180
 Db 121 VLVSQSGYLFIOTDKTIYTPGSTVLRYRIFTVNHKLLPVGRVTMWNINENPEGIPVKQDSL 180
 Qy 181 SSQNQLGVLPLSWDIPELVNMGMQKIRAYENSPOQVFSTEEFVEKYYLVPSEFVIEPTE 240
 Db 181 SSQNQLGVLPLSWDIPELVNMGMQKIRAYENSPOQVFSTEEFVEKYYLVPSEFVIEPTE 240
 Qy 241 KEYIYINKEGLEVTITARFLYKKGVEGTAFVIFGIDGEBQRIQLSPESLKRPIEDSGSEV 300
 Db 241 KEYIYINKEGLEVTITARFLYKKGVEGTAFVIFGIDGEBQRIQLSPESLKRPIEDSGSEV 300
 Qy 301 VLSRKVLDDGVQNPRAEDLVGKSLVYSATVILHSGSDMVQAEERSGIPITVSPYQIHFTKT 360
 Db 301 VLSRKVLDDGVQNPRAEDLVGKSLVYSATVILHSGSDMVQAEERSGIPITVSPYQIHFTKT 360
 Qy 361 PKYFKGMPFDMVFTNPDGSPAYRVPVAVOGEDEVQSLTQGDGVAKLSINTHPSQKPL 420
 Db 361 PKYFKGMPFDMVFTNPDGSPAYRVPVAVOGEDEVQSLTQGDGVAKLSINTHPSQKPL 420
 Qy 421 SITVRTKKQELSEAEQATRTMQALPYSTVGNSSNNYLHLSVLRTTELRLPGETLVNPLLEMD 480
 Db 421 SITVRTKKQELSEAEQATRTMQALPYSTVGNSSNNYLHLSVLRTTELRLPGETLVNPLLEMD 480
 Qy 481 RAHEAKIRYYTYLIMNKGRLKAGQVRPGQDLVVLPLSITTDPIPSFRIVAYYTLGA 540
 Db 481 RAHEAKIRYYTYLIMNKGRLKAGQVRPGQDLVVLPLSITTDPIPSFRIVAYYTLGA 540
 Qy 541 SGQREVADSVWVDVKDSCVGLVVKSGQSEDRQVPQGMTKITEGDHGARVVLVAVDK 600
 Db 541 SGQREVADSVWVDVKDSCVGLVVKSGQSEDRQVPQGMTKITEGDHGARVVLVAVDK 600
 Qy 601 GFVFLANKKLTQSKLWDVVEKADIGCTPGSGKDVAGVFS DAGLTFTSSSQQAQRAEL 660
 Db 601 GFVFLANKKLTQSKLWDVVEKADIGCTPGSGKDVAGVFS DAGLTFTSSSQQAQRAEL 660
 Qy 661 QCPQPAARRRRSVOLTEKMDKVGPXELRKCCEBDGMRNPMRPSCORRTFRFISLGEAC 720
 Db 661 QCPQPAARRRRSVOLTEKMDKVGPXELRKCCEBDGMRNPMRPSCORRTFRFISLGEAC 720
 Qy 721 KKVFLDCCNYIELRQARASHGLARSLNLDIEDIIAENIVSRSEPPESMLNWNEDLKE 780
 Db 721 KKVFLDCCNYIELRQARASHGLARSLNLDIEDIIAENIVSRSEPPESMLNWNEDLKE 780
 Qy 781 PPKNGISTKLMMIFLKDSITTWELIIVSNMSDKKGCIVADPFEVTVMQDFFIDLRLPSV 840
 Db 781 PPKNGISTKLMMIFLKDSITTWELIIVSNMSDKKGCIVADPFEVTVMQDFFIDLRLPSV 840
 Qy 841 RNEQVEIRAVLYNRQNELKVRVLELHNPAFCSLATTKRHHQQTITIPPKSSLSVPYVI 900
 Db 841 RNEQVEIRAVLYNRQNELKVRVLELHNPAFCSLATTKRHHQQTITIPPKSSLSVPYVI 900
 Qy 901 VPLKTGLQEVVEKAAVYHHFISDGVKSLKVYVEGIRMNKTAVRTLDPERLREGVQKE 960
 Db 901 VPLKTGLQEVVEKAAVYHHFISDGVKSLKVYVEGIRMNKTAVRTLDPERLREGVQKE 960
 Qy 961 DIPPADLSQVDPDTSSETRILLQGPVQAMTDEDAVDAERLKLHIVTSPGCGEONMIGWP 1020
 Db 961 DIPPADLSQVDPDTSSETRILLQGPVQAMTDEDAVDAERLKLHIVTSPGCGEONMIGWP 1020
 Qy 1021 TVIAVHYLDETQWKEKFGLEKQGALELIKKGYYTQQLAFROPSSAFAFVKRAPESTWLTA 1080
 Db 1021 TVIAVHYLDETQWKEKFGLEKQGALELIKKGYYTQQLAFROPSSAFAFVKRAPESTWLTA 1080
 Qy 1081 YVVKVPSLAVNLIAIDSOVLGAVKWLILEKQKDPGVFOEDAPVTHOBMIGLNNNEKD 1140
 Db 1081 YVVKVPSLAVNLIAIDSOVLGAVKWLILEKQKDPGVFOEDAPVTHOBMIGLNNNEKD 1140
 Qy 1141 MALTAFLVLSLOBAKDI CEEQVNSLPQSIITKAGDFLEANYMNLQSRYSYTAIAGYALQMG 1200
 Db 1141 MALTAFLVLSLOBAKDI CEEQVNSLPQSIITKAGDFLEANYMNLQSRYSYTAIAGYALQMG 1200

QY 1201 RLKGPLINKELTTAKQWNEDEPGKOLYNVENTSYALLALLQLKDFDVPVWVWNEQR 1260
DB 1201 RLKGPLINKELTTAKQWNEDEPGKOLYNVENTSYALLALLQLKDFDVPVWVWNEQR 1260
QY 1261 YGGGYGSTOATFWFQALAQYKQAPDHOELNLDVSLQLPSRSKITHRIHWESASLLR 1320
DB 1261 YGGGYGSTOATFWFQALAQYKQAPDHOELNLDVSLQLPSRSKITHRIHWESASLLX 1320
QY 1321 SEETKENEGFTVABKGOGTISVVTWYHAKADQLTCKNFDLKVTIKPAPETEKRPQDA 1380
DB 1321 SEETKENEGFTVABKGOGTISVVTWYHAKADQLTCKNFDLKVTIKPAPETEKRPQDA 1380
QY 1381 KNTMILEICTRYRGDQDATMSILDISMNTGAPDPTDDLKOLANGVDRIYSKYELDKAFSD 1440
DB 1381 KNTMILEICTRYRGDQDATMSILDISMNTGAPDPTDDLKOLANGVDRIYSKYELDKAFSD 1440
QY 1441 RNTLIYYLDKVSHSEDDCLAFKVHOFNVVELIQGAVKYAYYNLEESCTRFYHPEKEDG 1500
DB 1441 RNTLIYYLDKVSHSEDDCLAFKVHOFNVVELIQGAVKYAYYNLEESCTRFYHPEKEDG 1500
QY 1501 KLKLCRDELCRCABENCFIQKSDDKVTLLEERLDKACBPGVDYVYKTLVKVQLSNDPDE 1560
DB 1501 KLKLCRDELCRCABENCFIQKSDDKVTLLEERLDKACBPGVDYVYKTLVKVQLSNDPDE 1560
QY 1561 YIWAIEOTIKSGSDEVQVQOQRTFISPIKCREALKLEEKHYLMWGLSSDFWGEKPNLSY 1620
DB 1561 YIWAIEOTIKSGSDEVQVQOQRTFISPIKCREALKLEEKHYLMWGLSSDFWGEKPNLSY 1620
QY 1621 IIGKDTWVHWPDEECQDEENQKQCDLGAFTESMVVFGCPN 1663
DB 1621 IIGKDTWVHWPDEECQDEENQKQCDLGAFTESMVVFGCPN 1663

Search completed: December 22, 2004, 00:33:06
Job time : 205 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2004, 00:33:14 ; Search time 191 Seconds
(without alignments)
3115.599 Million cell updates/sec

Title: US-09-875-519A-22
Perfect score: 8609
Sequence: 1 MGPTSGPSLLLLLTHPLA.....KQCQDLGARTSVVFGCPN 1663

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1589859 seqs, 357834939 residues

Total number of hits satisfying chosen parameters: 1589859

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Sequence 6, Appli
Sequence 80, Appl
Sequence 115, Appl
Sequence 70, Appl
Sequence 2, Appli
Sequence 59, Appl
Sequence 7, Appli
Sequence 8, Appli
Sequence 9, Appli
Sequence 29, Appl
Sequence 30, Appl
Sequence 13, Appl
Sequence 59, Appl
Sequence 11, Appl
Sequence 81, Appl
Sequence 116, Appl
Sequence 17, Appl
Sequence 71, Appl
Sequence 15, Appl
Sequence 9, Appli
Sequence 5, Appli
Sequence 66, Appl
Sequence 5, Appli
Sequence 10, Appl
Sequence 12, Appl
Sequence 13, Appl
Sequence 405, Appl
Sequence 38, Appl
Sequence 10, Appl

US-09-875-519A-22
; Sequence 22, Application US/09875519A
; Patent No. US20020068059A1
; GENERAL INFORMATION:
; APPLICANT: Faries, Timothy C.
; TITLE OF INVENTION: Down-Regulation Resistant C3 Convertase
; FILE REFERENCE: 4-30443/A/IMU/PCT
; CURRENT APPLICATION NUMBER: US/09/875,519A
; CURRENT FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: PCT/GB97/00603
; PRIOR FILING DATE: 1997-03-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 22
; LENGTH: 1663
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-875-519A-22

ALIGNMENTS

RESULT 1
; Sequence 22, Application US/09875519A
; Patent No. US20020068059A1
; GENERAL INFORMATION:
; APPLICANT: Faries, Timothy C.
; TITLE OF INVENTION: Down-Regulation Resistant C3 Convertase
; FILE REFERENCE: 4-30443/A/IMU/PCT
; CURRENT APPLICATION NUMBER: US/09/875,519A
; CURRENT FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: PCT/GB97/00603
; PRIOR FILING DATE: 1997-03-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 22
; LENGTH: 1663
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-875-519A-22

Query Match 100.0%; Score 8609; DB 9; Length 1663;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MGPTSGPSLLLLLTHPLALGSPMYSITPNILRESEETWVLEAHDAQGDVFTVTVH 60
Db 1 MGPTSGPSLLLLLTHPLALGSPMYSITPNILRESEETWVLEAHDAQGDVFTVTVH 60

Oy 61 DFGPKLVLSSEKTVLTPATNMGNVTITIPANREFKSEKGRNFVTQATFGQVVEKV 120
Db 61 DFGPKLVLSSEKTVLTPATNMGNVTITIPANREFKSEKGRNFVTQATFGQVVEKV 120

Oy 121 VLVSLSQGVLFITQDXTIYTPGSTVLYRIFTVNHKLLPVGRTVMVNIENPESGIVKQDSL 180
Db 121 VLVSLSQGVLFITQDXTIYTPGSTVLYRIFTVNHKLLPVGRTVMVNIENPESGIVKQDSL 180

181 SSQNLGVLPLSDIPVLNMGQWKIRAYYENSPOQVFSTEPVEVKEYVLPSEFVIVPTE 240
Db 181 SSQNLGVLPLSDIPVLNMGQWKIRAYYENSPOQVFSTEPVEVKEYVLPSEFVIVPTE 240
QY 241 KFYIYNEKLEVTITARFLYGGKVEGTAFVIGTQGEORISLPSKRIPIEDGSGEV 300
Db 241 KFYIYNEKLEVTITARFLYGGKVEGTAFVIGTQGEORISLPSKRIPIEDGSGEV 300
QY 301 VLSRKVLDDGQNPRAEDLVGKSLYVSATVILHSGSDMVQABRSQIPVTSPIQIHFETK 360
Db 301 VLSRKVLDDGQNPRAEDLVGKSLYVSATVILHSGSDMVQABRSQIPVTSPIQIHFETK 360
QY 361 PKYFKGMPDLVAVFVTPNPGSPAYRVPVAVQGEDTVQSLTQDGVAKLSINTHPSQKPL 420
Db 361 PKYFKGMPDLVAVFVTPNPGSPAYRVPVAVQGEDTVQSLTQDGVAKLSINTHPSQKPL 420
QY 421 SITVTRTKQBSLSEAEQATRTMOALPYSTVGNSSNNYLHLVSLTELPGETLNVLNPLRMD 480
Db 421 SITVTRTKQBSLSEAEQATRTMOALPYSTVGNSSNNYLHLVSLTELPGETLNVLNPLRMD 480
QY 481 RAHEAKIRYTYILMKGRLKAGROVRBPQDGLVPLSITTDPIPSFRLVAVYTLIGA 540
Db 481 RAHEAKIRYTYILMKGRLKAGROVRBPQDGLVPLSITTDPIPSFRLVAVYTLIGA 540
QY 541 SQREVVADSVWVDKSDCVGSLVWKSQSEDRQVPQGMPLKTEGHDGARVVLVAVDK 600
Db 541 SQREVVADSVWVDKSDCVGSLVWKSQSEDRQVPQGMPLKTEGHDGARVVLVAVDK 600
QY 601 GVFLNKKNLTKOSKIWDVVEKADICTPGSGKDVAGVFSADGLTFTSSSQOQTAQRAEL 660
Db 601 GVFLNKKNLTKOSKIWDVVEKADICTPGSGKDVAGVFSADGLTFTSSSQOQTAQRAEL 660
QY 661 OCPOPAARRRSVQLETKMDKVGKPKELRKCCEGDMRENPMRFSQORRTFTSLGEC 720
Db 661 OCPOPAARRRSVQLETKMDKVGKPKELRKCCEGDMRENPMRFSQORRTFTSLGEC 720
QY 721 KKVFLDCNVTILRRORHARASHGLARSLNLDDEIIAENIVSRSEFFESMLWNVDELKE 780
Db 721 KKVFLDCNVTILRRORHARASHGLARSLNLDDEIIAENIVSRSEFFESMLWNVDELKE 780
QY 781 PKKNGISTKMLNFKLSDITTWELAVMSDKKGCIVADPPVTVMQDFFIDLPLPSV 840
Db 781 PKKNGISTKMLNFKLSDITTWELAVMSDKKGCIVADPPVTVMQDFFIDLPLPSV 840
QY 841 RNEQVEIRAVLYNQROQLKRVVLLHNPFCSLATTKRHHQOQITIPPKSSLSVPYVI 900
Db 841 RNEQVEIRAVLYNQROQLKRVVLLHNPFCSLATTKRHHQOQITIPPKSSLSVPYVI 900
QY 901 VPLATGLOEVEVKAAYVHHFISGVRSKLVVPEGRVXNKTVAVTLPDELREGVQKE 960
Db 901 VPLATGLOEVEVKAAYVHHFISGVRSKLVVPEGRVXNKTVAVTLPDELREGVQKE 960
QY 961 DIPPADLSQVPDTESETRILLQGTVAQMTEDAVDAERLKLIVTPSGCGEQNMIGWP 1020
Db 961 DIPPADLSQVPDTESETRILLQGTVAQMTEDAVDAERLKLIVTPSGCGEQNMIGWP 1020
QY 1021 TVIAVHVLDETEQWKEGKRGALRELKKGVTQOLAPROPSSAFAFVRASTWLTA 1080
Db 1021 TVIAVHVLDETEQWKEGKRGALRELKKGVTQOLAPROPSSAFAFVRASTWLTA 1080
QY 1081 YVVKVFLAVNLIAIDSQLCGAVKWLILEKQKPGVFOEDAPVTHQEMIGGLANNNEKD 1140
Db 1081 YVVKVFLAVNLIAIDSQLCGAVKWLILEKQKPGVFOEDAPVTHQEMIGGLANNNEKD 1140
QY 1141 MALTAFLVLSIQEAKDICEQVNSLRGSLKAGDFLEANYMLQRSYTVAGVALAQMG 1200
Db 1141 MALTAFLVLSIQEAKDICEQVNSLRGSLKAGDFLEANYMLQRSYTVAGVALAQMG 1200
QY 1201 RLKGPLNKLFTTAKDKRWEDSGKQLYNVVATSYALLALLOKDFDFVPPVWLNENQR 1260
Db 1201 RLKGPLNKLFTTAKDKRWEDSGKQLYNVVATSYALLALLOKDFDFVPPVWLNENQR 1260

1261 YYGSGYSTQATFVVFQALAQYQKQADPQHQLNLDVSLQPSRSSKITHRIHWESASLLR 1320
Db 1261 YYGSGYSTQATFVVFQALAQYQKQADPQHQLNLDVSLQPSRSSKITHRIHWESASLLR 1320
QY 1321 SSETKEGEGFTVTAEGKGQGTLSVWVTHYAKAKQDLTKNFKLVTKIPAPETEKRPQDA 1380
Db 1321 SSETKEGEGFTVTAEGKGQGTLSVWVTHYAKAKQDLTKNFKLVTKIPAPETEKRPQDA 1380
QY 1381 KNTWLIEICTRYRGDQATWSIILDISMTGFPADTDDKQLANGVDRYISKVELDKAFSD 1440
Db 1381 KNTWLIEICTRYRGDQATWSIILDISMTGFPADTDDKQLANGVDRYISKVELDKAFSD 1440
QY 1441 RNTLIILYLDKVSHEDEDCLAFKHQYFNVELIQPGAVKVYAYNLEESCTRFYHPEKEDG 1500
Db 1441 RNTLIILYLDKVSHEDEDCLAFKHQYFNVELIQPGAVKVYAYNLEESCTRFYHPEKEDG 1500
QY 1501 KNLKCRDELRCRBAENCFIQKSDDKVTLEERLDKACBEGVDYVYKTRLVKQLSNDDE 1560
Db 1501 KNLKCRDELRCRBAENCFIQKSDDKVTLEERLDKACBEGVDYVYKTRLVKQLSNDDE 1560
QY 1561 YVMAIEQTIKSGSDEVQVQGTFTFISPIKCRBALKLEEKHYLMGLSSDFWGEKPNLSY 1620
Db 1561 YVMAIEQTIKSGSDEVQVQGTFTFISPIKCRBALKLEEKHYLMGLSSDFWGEKPNLSY 1620
QY 1621 IIGKDTVWEHWPEDDEQDEENKQKQDILGAFTESMVVFGCPN 1663
Db 1621 IIGKDTVWEHWPEDDEQDEENKQKQDILGAFTESMVVFGCPN 1663

RESULT 2

US-09-842-758-41
; Sequence 41, Application US/09842758
; Publication No. US20030083244A1
; GENERAL INFORMATION:
; APPLICANT: Vernest, Corine A. M.
; APPLICANT: Fernandes, Elma R
; APPLICANT: Gerlach, Valerie
; APPLICANT: Shimkets, Richard A
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Soldog, Ferenc L
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Majumder, Kumud
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Radigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Burgess, Catherine E
; APPLICANT: Gangoli, Esha A
; APPLICANT: Smithson, Glenda
; APPLICANT: Rastelli, Luca
; APPLICANT: MacDougall, John R
; APPLICANT: Taupier, Raymond J
; APPLICANT: Grosse, William M
; APPLICANT: Edward, Szekeres S
; APPLICANT: Alsbrook II, John P
; TITLE OF INVENTION: No. US20030083244A1 Proteins and Nucleic Acids Encoding Same
; FIE REFERENCE: 15966-783
; CURRENT APPLICATION NUMBER: US/09/842,758
; CURRENT FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/200,158
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/200,513
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,780
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/201,006
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,007
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,236
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,238
; PRIOR FILING DATE: 2000-05-01

; PRIOR APPLICATION NUMBER: 60/201,186
 ; PRIOR FILING DATE: 2000-05-02
 ; PRIOR APPLICATION NUMBER: 60/201,474
 ; PRIOR FILING DATE: 2000-05-03
 ; PRIOR APPLICATION NUMBER: 60/201,508
 ; PRIOR FILING DATE: 2000-05-03
 ; PRIOR APPLICATION NUMBER: 60/220,591
 ; PRIOR FILING DATE: 2000-07-25
 ; PRIOR APPLICATION NUMBER: 60/232,678
 ; PRIOR FILING DATE: 2000-09-15
 ; PRIOR APPLICATION NUMBER: 60/263,217
 ; PRIOR FILING DATE: 2001-01-22
 ; PRIOR APPLICATION NUMBER: 60/265,160
 ; PRIOR FILING DATE: 2001-01-30
 ; NUMBER OF SEQ ID NOS: 113
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 41
 ; LENGTH: 1663
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-842-758-41

Query Match 99.9%; Score 8599; DB 10; Length 1663;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1661; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGPTSGPSLLLLTHLPALGSPMYSIITPNILRLESEETVLEAHDAQGDPVTVVH 60
 DB 1 MGPTSGPSLLLLTHLPALGSPMYSIITPNILRLESEETVLEAHDAQGDPVTVVH 60
 QY 61 DFPKGLVLSSEKTLVTPATNHGNTFTIPANREFKSEKGRNKFVTVQATFGTQWEKV 120
 DB 61 DFPKGLVLSSEKTLVTPATNHGNTFTIPANREFKSEKGRNKFVTVQATFGTQWEKV 120
 QY 121 VLVSQSGYLFICDQKTIYTPGSTVLYRIFTVNHKLLPVGRVTWNINPEGIPYKQDSL 180
 DB 121 VLVSQSGYLFICDQKTIYTPGSTVLYRIFTVNHKLLPVGRVTWNINPEGIPYKQDSL 180
 QY 181 SSQQLGVPLSLMDIPELVNMGMQKIRAYENSPQOVFTSEFEVKEVYLPSEV-VEPTE 240
 DB 181 SSQQLGVPLSLMDIPELVNMGMQKIRAYENSPQOVFTSEFEVKEVYLPSEV-VEPTE 240
 QY 241 KFYIYNEKGLVETIITARFLYGKVGRTAFVIGIOGQOR-SLPSLKRITIEGSGEV 300
 DB 241 KFYIYNEKGLVETIITARFLYGKVGRTAFVIGIOGQOR-SLPSLKRITIEGSGEV 300
 QY 301 VLSKVLLDGVQNPRAEDLVCKSLYVSATVILHSGSDMVQAESGIPITSPYQHFTKT 360
 DB 301 VLSKVLLDGVQNPRAEDLVCKSLYVSATVILHSGSDMVQAESGIPITSPYQHFTKT 360
 QY 361 PKYFKGMPDLMVFTNPDGSPAYRVPVAVQGEDTVQSITOGDGVAKLSINTHPSOKPL 420
 DB 361 PKYFKGMPDLMVFTNPDGSPAYRVPVAVQGEDTVQSITOGDGVAKLSINTHPSOKPL 420
 QY 421 SITVTRKKQELSAEQATRMQALPYSTVGNSNNYLHLSVLRLTELRPGSETLNVNLLRMD 480
 DB 421 SITVTRKKQELSAEQATRMQALPYSTVGNSNNYLHLSVLRLTELRPGSETLNVNLLRMD 480
 QY 481 RAHEAKIRYTYLIMNKGRLKAGQVRPBGDLVPLSITTDTPSPRLVAYYTLIGA 540
 DB 481 RAHEAKIRYTYLIMNKGRLKAGQVRPBGDLVPLSITTDTPSPRLVAYYTLIGA 540
 QY 541 SQGEVVDVSVVDKSDCVSGSLVWKSQSGEDRQVPVGGQMTLKIEGDHGAARVLVADK 600
 DB 541 SQGEVVDVSVVDKSDCVSGSLVWKSQSGEDRQVPVGGQMTLKIEGDHGAARVLVADK 600
 QY 601 GVFVNLKKNLKTOSKIWDVVEKADIGCTPGSGKDYAGVFSADAGLTFTSSGQCTQRAEL 660
 DB 601 GVFVNLKKNLKTOSKIWDVVEKADIGCTPGSGKDYAGVFSADAGLTFTSSGQCTQRAEL 660
 QY 661 QCPOPAARRRRSVQLETKRMDKVGKPKELRKCCEDGMRENPMRFSCQRRTRFISLGEAC 720
 DB 661 QCPOPAARRRRSVQLETKRMDKVGKPKELRKCCEDGMRENPMRFSCQRRTRFISLGEAC 720

QY 721 KKVFDCQNYITELRQCHASHLGLARSNLDEDLIAENIVSRSEFPESLWVDELDKE 780
 DB 721 KKVFDCQNYITELRQCHASHLGLARSNLDEDLIAENIVSRSEFPESLWVDELDKE 780
 QY 781 PPKNGISTKLNNIFLKDSITTWELLAVSMDSKKIGICVADPFEVTVNQOFFIDLRLPYSV 840
 DB 781 PPKNGISTKLNNIFLKDSITTWELLAVSMDSKKIGICVADPFEVTVNQOFFIDLRLPYSV 840
 QY 841 RNEQVEIRAVLYNRYRONQELKVRVELLHNPAPCSLATTKRHHOQTITTPPKSSLSVPVVI 900
 DB 841 RNEQVEIRAVLYNRYRONQELKVRVELLHNPAPCSLATTKRHHOQTITTPPKSSLSVPVVI 900
 QY 901 VPLKTGLQEVVEKAAVYHHFISDGVKRSKLVVPEGIRMNKTVAVRTLDPERLREGVQKE 960
 DB 901 VPLKTGLQEVVEKAAVYHHFISDGVKRSKLVVPEGIRMNKTVAVRTLDPERLREGVQKE 960
 QY 961 DIPADLSQVDPDTESETRILLQGTTPVAQMTEDAVDAERLKHILVTPSGCGEONMIGWTP 1020
 DB 961 DIPADLSQVDPDTESETRILLQGTTPVAQMTEDAVDAERLKHILVTPSGCGEONMIGWTP 1020
 QY 1021 TVIAVHYLDETEQWERKLEKQGALELIKKGYSITQOLAFROPSSAFAPVRAAPSTWUTA 1080
 DB 1021 TVIAVHYLDETEQWERKLEKQGALELIKKGYSITQOLAFROPSSAFAPVRAAPSTWUTA 1080
 QY 1081 YVVKVFSLAUNLIAIDSOVLGAVKWLILLEKQKPGVFCEDAPVTHQEMIIGGLRNNNEKD 1140
 DB 1081 YVVKVFSLAUNLIAIDSOVLGAVKWLILLEKQKPGVFCEDAPVTHQEMIIGGLRNNNEKD 1140
 QY 1141 MALTAFLVLSQBAKIDICEQVNSLPGSITKAGDFLEANYMNLORSYTVAIAGYALAQMG 1200
 DB 1141 MALTAFLVLSQBAKIDICEQVNSLPGSITKAGDFLEANYMNLORSYTVAIAGYALAQMG 1200
 QY 1201 RLKGPLLNKFLTTAKDKNRWEDPGKQLYNVEATSVALLALQLKDFDVPVVRWLNQOR 1260
 DB 1201 RLKGPLLNKFLTTAKDKNRWEDPGKQLYNVEATSVALLALQLKDFDVPVVRWLNQOR 1260
 QY 1261 YGGGYGSTQATPMVQALAQYKDAPDHOELNLDVLSQLPSRSSKITTHRIHWSASLLR 1320
 DB 1261 YGGGYGSTQATPMVQALAQYKDAPDHOELNLDVLSQLPSRSSKITTHRIHWSASLLR 1320
 QY 1321 SEETKNEGFVTVAGKGGOTLSVVMYHAKAKDQLTCKNPKLVYTIKPAETEKRPDA 1380
 DB 1321 SEETKNEGFVTVAGKGGOTLSVVMYHAKAKDQLTCKNPKLVYTIKPAETEKRPDA 1380
 QY 1381 KNTMILEICTRYRGDQDATMSILDISMMTGFPAPDTDDLKQLANGVDVRIYSKYELDKAFSD 1440
 DB 1381 KNTMILEICTRYRGDQDATMSILDISMMTGFPAPDTDDLKQLANGVDVRIYSKYELDKAFSD 1440
 QY 1441 RNTLIIYLDKVSHEDDCLAFKVHGVNVELLOPGAVKYAYNLEESCTRFYHPEKEDG 1500
 DB 1441 RNTLIIYLDKVSHEDDCLAFKVHGVNVELLOPGAVKYAYNLEESCTRFYHPEKEDG 1500
 QY 1501 KLNKLRDELRCABENCFIQKSDDKVLEERLDRACEPGVDYVYKTRLVKVLNSNDFDE 1560
 DB 1501 KLNKLRDELRCABENCFIQKSDDKVLEERLDRACEPGVDYVYKTRLVKVLNSNDFDE 1560
 QY 1561 YIMAIQTITKSGSDVQVQOQRTFISPIKCRALALEKCHYLMGLSSDFWGEKPNLSY 1620
 DB 1561 YIMAIQTITKSGSDVQVQOQRTFISPIKCRALALEKCHYLMGLSSDFWGEKPNLSY 1620
 QY 1621 IIGKDTWHEHWPDEECQDEENKQCCQDLGAFTESMVFPGCPN 1663
 DB 1621 IIGKDTWHEHWPDEECQDEENKQCCQDLGAFTESMVFPGCPN 1663

RESULT 3
 US-10-379-747-2
 ; Sequence 2, Application US/10379747
 ; Publication No. US20040023874A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Burgess, Catherine E.;
 ; APPLICANT: Chant, John S.;


```

RESULT 4
US-10-174-333-41
; Sequence 41, Application US/10174333
; Publication No. US2004029220A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Zethusen, Bryan D.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Majumder, Kumud
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Smithson, Glenda
; APPLICANT: Rastelli, Luca
; APPLICANT: MacDougall, John R.
; APPLICANT: Taupier, Raymond J.
; APPLICANT: Grosse, William M.
; APPLICANT: Szekeres, Edward S.
; APPLICANT: Alsobrook, John P.
; APPLICANT: Anderson, David W.
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Li, Li Mei
; APPLICANT: Zhong,
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-783 C1P1
; CURRENT APPLICATION NUMBER: US/10/174,333
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: 60/193,664
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 60/194,614
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,063
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,066
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,067
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,068
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,069
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,070
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,510
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/219,855
; PRIOR FILING DATE: 2000-07-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: Curasequid version 0.1
; SEQ ID NO 41
; LENGTH: 1663
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-174-333-41

Query Match 99.9%; Score 8596; DB 15; Length 1663;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1661; Conservative 1; Mismatches 0; Gaps 0;

QY 1 MGPTSGSLLLLTHLPLALGSPMSIITPNILRLESETWVLEAHDAGQDPVTVTH 60
DB 1 MGPTSGSLLLLTHLPLALGSPMSIITPNILRLESETWVLEAHDAGQDPVTVTH 60
QY 61 DFPCKKLVLSSEKIVLTPTATNMGNTFTIPANREFKSEKGRNKFVTVOATFGTQWEKV 120

```

QY 1201 RLKGLPLNKLFTAKDKNRWEPGKOLYNVEATSVALLALLQLKDFDFVPPVWVWLNQOR 1260
DB 1201 RLKGLPLNKLFTAKDKNRWEPGKOLYNVEATSVALLALLQLKDFDFVPPVWVWLNQOR 1260
QY 1261 YGCGYGSQATPMWFQALAQKQAPDHOELNVLQPSRSKITHRIHWESASLLR 1320
DB 1261 YGCGYGSQATPMWFQALAQKQAPDHOELNVLQPSRSKITHRIHWESASLLR 1320
QY 1321 SEETKENEGFTVTAEGKGQGTLSVVMYHAKAKDQLTCKNPKLVTKIAPAPETEKRPQDA 1380
DB 1321 SEETKENEGFTVTAEGKGQGTLSVVMYHAKAKDQLTCKNPKLVTKIAPAPETEKRPQDA 1380
QY 1381 KNTMILIECTRYRGDQDATMSILDLSMTGFPAPDITDLDKQLANGVDRVYSIKYELDKAFSD 1440
DB 1381 KNTMILIECTRYRGDQDATMSILDLSMTGFPAPDITDLDKQLANGVDRVYSIKYELDKAFSD 1440
QY 1441 RNTLIYLDKVSSEDDCLAFKHQYFNVVELIOPGAVKYAYYNLEESCTRFYHPEKEDG 1500
DB 1441 RNTLIYLDKVSSEDDCLAFKHQYFNVVELIOPGAVKYAYYNLEESCTRFYHPEKEDG 1500
QY 1501 KLNKLCRDELCRCABENCFIQSDDKVTLLEERLDKACBPGVDVYVTKLKVQLSNDFDE 1560
DB 1501 KLNKLCRDELCRCABENCFIQSDDKVTLLEERLDKACBPGVDVYVTKLKVQLSNDFDE 1560
QY 1561 YIMAEOTIKSGSDEYVQVQQRITFISPIKCREALKLEEKHYLMWGLSSDFWGEKPNLSY 1620
DB 1561 YIMAEOTIKSGSDEYVQVQQRITFISPIKCREALKLEEKHYLMWGLSSDFWGEKPNLSY 1620
QY 1621 IIGKOTVWEHPBEDCODENQKQCODLIGATESWVFGCPN 1663
DB 1621 IIGKOTVWEHPBEDCODENQKQCODLIGATESWVFGCPN 1663

RESULT 5

US-09-842-758-43
; Sequence 43, Application US/09842758
; Publication No. US20030083244A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Malvankar, Uriel M.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Zernusen, Bryan D.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Majumder, Rumud
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Smithson, Glennda
; APPLICANT: Kastelli, Luca
; APPLICANT: MacDougall, John R.
; APPLICANT: Taupier, Raymond J.
; APPLICANT: Grosse, William M.
; APPLICANT: Edward, Szekeres S.
; APPLICANT: Alsobrook II, John P.
; TITLE OF INVENTION: No. US20030083244A1e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-783
; CURRENT APPLICATION NUMBER: US/09/842, 758
; CURRENT FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/200,158
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/200,613
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,780
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/201,006
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,007

; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,236
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,236
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,186
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: 60/201,474
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/201,508
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/220,591
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/232,678
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 60/263,217
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 60/265,160
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 1663
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-842-758-43

Query Match 78.4%; Score 6752; DB 10; Length 1663;
Best Local Similarity 77.1%; Pred. No. 0;
Matches 1485; Conservative 178; Mismatches 197; Indels 6; Gaps 5;

QY 1 MGPTSGPSEL--LILLLLTHPLALGSPMYSIITNIIIRLSEETMVLSEAHDAQDQVPVTVT 58
DB 1 MGPASGSQLVLLVLLASSPLALGIPMYSIITNIIIRLSEETMVLSEAHDAQDQDIPVTVT 60
QY 59 VHDPEGKLVLSSEKTVLTATNMGNVTFTIPANREFKSEKRNKVFVVOATFGTVVE 118
DB 61 VQDFL-KRQVLTSEKTVLTGASGHLRSVSIKIPASKEFNSDKGKHVTVVWAFGETVVE 119
QY 119 KVLVLSQSGYLFITDRTIYTPGSTVLYRIFTNNHKLPLVGTNNVNIENEGIPVKQD 178
DB 120 KAVVVSFQSGYLFITDRTIYTPGSTVLYRIFTNNHKLPLVGTNNVNIENEGIPVKQD 179
QY 179 SLSSONQLVPLSDIPELVNMGMOWIRAYYNSPOQVFSTFEVEKVVLPSPFVIVP 238
DB 180 ILSSNNQGLPLSWNIPELVNMGMOWIRAFYEHAPQIFSAFEFVEYVLPSPFVIVP 239
QY 239 TEKPYIYNEKGLVTTIARFLYKQKVEGTAFVIFGIQDGEQRIISLPESLKRIPIEDGSG 298
DB 240 TETFYIIDDPNGLVSIIAKFLYKGNVDGTAFVIFGVQDGGKKISLAHSLTRVIEDGVG 299
QY 299 EVLRSKVLVDGONPRAEDLVGKSLVSVATVILHSGSDMVQAERSGIPVITSPYQIHT 358
DB 300 DAVLTRKVLMEGVPSNADALVGKSLVSVTVILHSGSDMVQAERSGIPVITSPYQIHT 359
QY 359 KIPKYFKGMPFDLMVFTVTPDGSPPAYRVAVQGEDTVQSLTQDGVAKLSINTHPSOK 418
DB 360 KIPKFKGMPFDLMVFTVTPDGSPPASKVLVVTQGSN-AKALTQDDGVAKLSINTHPSOK 418
QY 419 PLSTIVTRTKKQELSEAEQATRMALPYSTVGNSSNNYHLSTLRTELRTPEGTINNVFLIR 478
DB 419 PLSTIVTRTKKQELSEAEQATRMALPYSTVGNSSNNYHLSTLRTELRTPEGTINNVFLIR 478
QY 479 MDRAHEAKIRYTYLNNKGLLKAGQVREPODLVPLSTITTDIFISFLVAYTYLI 538
DB 479 TDPGHEAKIRYTYLNNKGLLKAGQVREPODLVPLSTITTDIFISFLVAYTYLI 538
QY 539 GASGQREVVDVWVDVKDCVSGSLVVKSGQSDRQVPCQQTMLXIEGHGAVRVLVAV 598
DB 539 GASGQREVVDVWVDVKDCVSGSLVVKSGQSDRQVPCQQTMLXIEGHGAVRVLVAV 597
QY 599 DXGVFLNKKKLTQSKINDWVEKADIGCTPGSGKDYAGVFSAGLTFTSSSSQQTAAQA 658
DB 599 DXGVFLNKKKLTQSKINDWVEKADIGCTPGSGKDYAGVFSAGLTFTSSSSQQTAAQA 658

598 DKGVFVNKKKNTQSKIMDVWEKADIGCTPGSGKNYAGVEMDAGLAFKTSOGLQTEQRA 657
659 ELQCFQPAARRRSVQTERMBKVGKY-PKELRKCCEGDMENPMRSCORTRFISLG 717
658 DLECTKPAARRRSVQMERMDKAGYTDGURKCCEDMDIPMRISQRRALITQG 717
718 EACKVFLDCNVITELRQARASHLGLARSNLDEIIAENIVSRSEFFPESWLNVED 777
718 ENCIFADICNCHITKREQHRDHLGLARSELESDIIPEDIIISRSHPQSWLWITEE 777
778 LKEPKNGISTKLMNIFPKDSITTWELAVMSDKKICVADPEEVVMQDFIDLRUPY 837
778 LKEPKNGISTKMNIFPKDSITTWELAVMSDKKICVADPEEVVMQDFIDLRUPY 837
838 SVRNEQVEIRAVLYNQROELKVRVELLHNPFCSLATTKRHQQITIPPKSSLSVP 897
838 SVRNEQVEIRAVLYNQROELKVRVELLHNPFCSMATXNRYFQTIKIPKSSVAVP 897
898 YVIVPLKIGQVEVKAHVHHSISGVKSLKVPGEIRMKNTVAVRTLDPERLGRGV 957
898 YVIVPLKIGQVEVKAHVHHSISGVKSLKVPGEIRMKNTVAVRTLDPERLGRGV 957
958 QKEDIPADISDQVDPDESSTRILQSTPVAQMTEDAVDAERLKLHIVTPSGCGQNMIG 1017
958 QKEDIPADISDQVDPDESSTRILQSTPVAQMTEDAVDAERLKLHIVTPSGCGQNMIG 1017
1018 MTPVIAVHVDTEQWEKEKGLKRGALFKGTQQLAFROPSSAPFAVFKRAPSTW 1077
1018 MTPVIAVHVDTEQWEKEKGLKRGALFKGTQQLAFROPSSAPFAVFKRAPSTW 1077
1078 LTAVVVKFSLAVNLIAIDSOVLCGAVKMLILEKQKPDGVQFQEDAPVHQEMIGLNN 1137
1078 LTAVVVKFSLAVNLIAIDSOVLCGAVKMLILEKQKPDGVQFQEDAPVHQEMIGLNN 1137
1138 EKMALTAFLVLSIQEAKDICEQVNSLPGSITKAGDFLEANYMNLQRSYTVIAIAGYALA 1197
1138 EADVSLTAFLVLSIQEAKDICEQVNSLPGSITKAGDFLEANYMNLQRSYTVIAIAGYALA 1197
1198 QMGLKGLPLNKFTTTAKDNREDPKQLYNNVATSYALLALLOKDFPVPVVRWLN 1257
1198 LMKLEBPYLGKFLNTAKDNREDPKQLYNNVATSYALLALLOKDFPVPVVRWLN 1257
1258 EORYGGYGGSTQATFVFOALAQYQKADPDHQLNLDLSLOLPSRSSKITHRIWESAS 1317
1258 EORYGGYGGSTQATFVFOALAQYQKADPDHQLNLDLSLOLPSRSSKITHRIWESAS 1317
1318 LLRSEETKENGFTVTAEGKGQGLSVVWYHAKAQDLTKNFKDLKVTIKPAPETERKP 1377
1318 LLRSEETKENGFTVTAEGKGQGLSVVWYHAKAQDLTKNFKDLKVTIKPAPETERKP 1377
1378 QDAKNTMILEICTRYRQDQATWSILDISMTGTFAPDQDLDKOLANGVDVRYISKYELDKA 1437
1378 EAKNTMILEICTRYRQDQATWSILDISMTGTFAPDQDLDKOLANGVDVRYISKYELDKA 1437
1438 FSDRNTLITLKVSHSEDDCLAFKHQYFNVELIQGAVKVVAYYNNLEESCTRFYHPEK 1497
1438 FSNKNTLITLKVSHSEDDCLAFKHQYFNVELIQGAVKVVAYYNNLEESCTRFYHPEK 1497
1498 EDGKLNKLDLRCRCAEENCFQKSDKVTLEERLDKACEPGVDVYVKTIRLVKQLSND 1557
1498 DDGMLSKLCHSEMCRCABENCFQKSDKVTLEERLDKACEPGVDVYVKTIRLVKQLSND 1557
1558 FDEYVIMABQITKSGSDEVQVQQRFTIPISKREALKLEEKHYLMWLSDFGKRPEN 1617
1558 FDEYVIMABQITKSGSDEVQVQQRFTIPISKREALKLEEKHYLMWLSDFGKRPEN 1617
1618 LSYIIGKDTVWEHWPEDQEDENOKQODLGAFTESMVVFCPN 1663
1618 LSYIIGKDTVWEHWPEDQEDENOKQODLGAFTESMVVFCPN 1663

RESULT 6

US-10-174-333-43

Sequence 43, Application US/10174333
Publication No. US20040029220A1
GENERAL INFORMATION:
APPLICANT: Vernet, Corine A.M.
APPLICANT: Fernandes, Elma R.
APPLICANT: Gerlach, Valerie
APPLICANT: Malyankar, Uriel M.
APPLICANT: Boldog, Ferenc L.
APPLICANT: Zerhusen, Bryan D.
APPLICANT: Spyttek, Kimberly A.
APPLICANT: Majumder, Kumud
APPLICANT: Tchernev, Velizar T.
APPLICANT: Padigaru, Muralidhara
APPLICANT: Patturajan, Meera
APPLICANT: Burgess, Catherine E.
APPLICANT: Gangoli, Esha A.
APPLICANT: Smithson, Glennda
APPLICANT: Rastelli, Luca
APPLICANT: MacDougall, John R.
APPLICANT: Taupier, Raymond J.
APPLICANT: Grosse, Edward S.
APPLICANT: Szekeres, Edward S.
APPLICANT: Alsobrook, John P.
APPLICANT: Anderson, David W.
APPLICANT: Guo, Xiaojia (Sasha)
APPLICANT: Li, Li
APPLICANT: Zhong, Mei
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 15966-783 CIPA
CURRENT APPLICATION NUMBER: US/10/174,333
CURRENT FILING DATE: 2002-06-18
PRIOR APPLICATION NUMBER: 60/193,664
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 60/194,614
PRIOR FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: 60/195,063
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/195,066
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/195,067
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/195,068
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/195,069
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/195,070
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/195,510
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/219,855
PRIOR FILING DATE: 2000-07-21
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 186
SOFTWARE: Curated version 0.1
SEQ ID NO 43
LENGTH: 1663
TYPE: PRT
ORGANISM: Mus musculus
US-10-174-333-43

Query Match 78.4%; Score 6752; DB 15; Length 1663;
Best Local Similarity 77.1%; Pred. No. 0;
Matches 1285; Conservative 178; Mismatches 197; Indels 6; Gaps 5;
Qy 1 MGTPSGPSL--LILLMIHLPLALGSPMYSIITPNIRLESEBTWVLEADAGQDVPVTVT 58
Db 1 MGTPSGPSL--LILLMIHLPLALGSPMYSIITPNIRLESEBTWVLEADAGQDVPVTVT 60
Qy 59 VHPDPKLVLSSEKTVLTPAHMGNVTFIPANREFSKGNKFTVVOATGTOVVE 118
Db 61 VQDFL-KRQVLTSEKTVLTGASGHLRSVSIKIPASKEFNSDKGHKYVTVVANFGETVVE 119
Qy 119 KVVLSVLSQSYLFIQTDKTIYTPGSTVLYRIFTVNHKLLPVGRVTVMNIENPEGIPVKQD 178

Db 120 KAVVBSFQSGFIQIDKTIYIPGTVLRYIFTVDNNLLPVGKTVWLLIETPDGIPVKRD 179
Qy 179 SLSSQNLGVLPISWDIPBLVNMGNWKIRAYIENSPOQVFTSEFEVKEYVLPSPSEVIVEP 238
Db 180 ILSNNQHGIPLSNWIPBLVNMGNWKIRAFYEHAPKQIFSAPEVKEYVLPSPSEVIVEP 239
Qy 239 TEFYIYIYNEKGLLEVITAFRLYGLKVEGTAFVIFIGIQDGEORSLPESIKRIPIEDGSG 298
Db 240 TEFYIYIDDPNGLEVISIAKELYGNVDGTAFFVIFGVQDQDKISLAHSLTRVVIDGVG 299
Qy 299 EVVLSKVLVDGONPRAEDILVKSILYSATVILHSGSDMVQABRSIGPIVITSYQIHFT 358
Db 300 DAVLTRKVLMEGRPNADALVKSILYSATVILHSGSDMVQABRSIGPIVITSYQIHFT 359
Qy 359 KTEPKFKPMPFDMVFNPDGSPAYRVPVAVQGEDTVQSITQDGVAKLSINTHPSQK 418
Db 360 KTEPKFKPMPFDMVFNPDGSPASKVLVVTQGSN-AKALTQDDGVAKLSINTHPSRQ 418
Qy 419 PLSITVTRKQELSEABQATRTMOALPSTVGNSSNNYLHLSVRLTELAPGETLNVNLLR 478
Db 419 PLTITVTRKQELSEABQATRTMOALPSTVGNSSNNYLHLSVRLTELAPGETLNVNLLR 478
Qy 479 MDRAHAKIRYIYTLIMNKGRLKAGQVREPGQDLVLSITDFTIPSPRLVAYYTLI 538
Db 479 TDPGHEAKIRYIYTLIMNKGRLKAGQVREPGQDLVLSITDFTIPSPRLVAYYTLI 538
Qy 539 GASQREVADSVWVDVKDSCVGLSVKSGQSEDQVPVPGQOMLKTREGHGAVLVAV 598
Db 539 GASQREVADSVWVDVKDSCVGLSVKSGQSEDQVPVPGQOMLKTREGHGAVLVAV 597
Qy 599 DKGVFVNLKKNLTQSKIWDVVEKADIGCTPGSGKDYAGVFSADAGLTFTSSSGQOQA 658
Db 599 DKGVFVNLKKNLTQSKIWDVVEKADIGCTPGSGKDYAGVFSADAGLTFTSSSGQOQA 657
Qy 659 ELQCPQAPARRRSVOLTERRDKVGYK-PRELKCCEDGMRENPMRFSQORRFPFSLG 717
Db 659 ELQCPQAPARRRSVOLTERRDKVGYK-PRELKCCEDGMRENPMRFSQORRFPFSLG 717
Qy 718 EACKVFLDCNVTILRLQHARASHGLGARSNDELIIABENIVSRSEPPESWLMNVED 777
Db 718 ENCIKAFIDCCNHTKLRQHRDRHVLGLARSELEEDIIPEEDIISRHPPQSWLWTEE 777
Qy 778 LKSPKNGISTKLMNIFLKDSTITWEILLAVSNDKKGICVADDPREVWQDFFIDRLPY 837
Db 778 LKSPKNGISTKLMNIFLKDSTITWEILLAVSNDKKGICVADDPREVWQDFFIDRLPY 837
Qy 838 SVVRNEQVEIRAVLYNYRQNELKVRVELLHNPFCSLATTKRRHQQTITIPKSSLSVP 897
Db 838 SVVRNEQVEIRAVLYNYRQNELKVRVELLHNPFCSLATTKRRHQQTITIPKSSLSVP 897
Qy 898 YVAVBLKTGLQVEVKAAYHYHFI SDGVRKSLKVVPGEIRMNKTVAVRITLDPBLGREGV 957
Db 898 YVAVBLKTGLQVEVKAAYHYHFI SDGVRKSLKVVPGEIRMNKTVAVRITLDPBLGREGV 957
Qy 958 QKEDIPADISQVPTSETRIILQGTPTVPAQMTEDAVDAERLKHILVTPSGCGEQNMIG 1017
Db 958 QKEDIPADISQVPTSETRIILQGTPTVPAQMTEDAVDAERLKHILVTPSGCGEQNMIG 1017
Qy 1018 MTETVAVHVLDETEQWEGLEKQCGALELTKKGYTOOLAFROPSSAFAPVREASTW 1077
Db 1018 MTETVAVHVLDETEQWEGLEKQCGALELTKKGYTOOLAFROPSSAFAPVREASTW 1077
Qy 1078 LTAYVVKVSLAVNLIAIDSVLCGAVKWLILEKQKPGVQFQEDAPVTHQEMIGLNN 1137
Db 1078 LTAYVVKVSLAVNLIAIDSVLCGAVKWLILEKQKPGVQFQEDAPVTHQEMIGLNN 1137
Qy 1138 EKDMALTAFLVLSIQAKDICEOVNSLPGSTTKAGDFLEANYMQLORSYVATAGYALA 1197
Db 1138 EKDMALTAFLVLSIQAKDICEOVNSLPGSTTKAGDFLEANYMQLORSYVATAGYALA 1197
Qy 1198 QMGLKGPPLNKFLTAKDKNRWEDPGKQLYNVEATSYALLALILQKDFDVPVVRWLN 1257
Db 1198 QMGLKGPPLNKFLTAKDKNRWEDPGKQLYNVEATSYALLALILQKDFDVPVVRWLN 1257

Db 1198 LMNKLSEPYLKPNTAKDRNWEEDPOQLYNVEATSYALLALILQKDFDVPVVRWLN 1257
Qy 1258 EORYGGYGGYSTOATFMVFOALAQYOKDAPHOELNLDVSLQPSRSSKITHRIHWESAS 1317
Db 1258 EORYGGYGGYSTOATFMVFOALAQYQIDVDPDKDNDVSHFUPSRSSATTFRLLWENGN 1317
Qy 1318 LLRSEETKNEBGTAVTAEGRGQGTLSVVTVMYHAKAKDQJTONKFDLKVTIKPAPETERP 1377
Db 1318 LLRSEETKNEBGTAVTAEGRGQGTLSVVTVMYHAKAKDQJTONKFDLKVTIKPAPETERP 1377
Qy 1378 QDANKWILECTRYRGDQDATMSILDISMVTFGAPDQDOLKOLANGVDVRYISKVELDKA 1437
Db 1378 QDANKWILECTRYRGDQDATMSILDISMVTFGAPDQDOLKOLANGVDVRYISKVELDKA 1437
Qy 1438 FSDRNTLIYLDKSHSEDDCLAFKVHQPYNVELIQGAVKVYAYNLEESCTRFYHPEK 1497
Db 1438 FSDRNTLIYLDKSHSEDDCLAFKVHQPYNVELIQGAVKVYAYNLEESCTRFYHPEK 1497
Qy 1498 EDGKLNKLCRDELRCRAENCFIQKSDDKVTLLEERLDKACEPGVDVYKTRLVKQLSND 1557
Db 1498 EDGKLNKLCRDELRCRAENCFIQKSDDKVTLLEERLDKACEPGVDVYKTRLVKQLSND 1557
Qy 1558 FDEYMAIEQCTIKSGSDEVQVQOQRTFISFIKCREALKLEEKHYLMGLSSDFWGEKPN 1617
Db 1558 FDEYMAIEQCTIKSGSDEVQVQOQRTFISFIKCREALKLEEKHYLMGLSSDFWGEKPN 1617
Qy 1618 LSYIIGKDTWVYHWPREDCEQDEENOKGODLGAFTESVVFCCPN 1663
Db 1618 LSYIIGKDTWVYHWPREDCEQDEENOKGODLGAFTESVVFCCPN 1663

RESULT 7

US-09-842-758-42
Sequence 42, Application US/09842758
Publication No. US20030083244A1
GENERAL INFORMATION:
APPLICANT: Vernet, Corine A. M.
APPLICANT: Fernandes, Elma R
APPLICANT: Gerlach, Valerie
APPLICANT: Shimkete, Richard A
APPLICANT: Malyankar, Uriel M
APPLICANT: Boldog, Ferenc L
APPLICANT: Zerhusen, Bryan D
APPLICANT: Spytek, Kimberly A
APPLICANT: Majumder, Kumud
APPLICANT: Tchernev, Velizar T
APPLICANT: Padigaru, Muralidhara
APPLICANT: Patturajan, Meera
APPLICANT: Burgess, Catherine E
APPLICANT: Gangoli, Esha A
APPLICANT: Smithson, Glennnda
APPLICANT: Rastelli, Luca
APPLICANT: MacDougall, John R
APPLICANT: Taupier, Raymond J
APPLICANT: Grosse, William M
APPLICANT: Edward, Szekeres S
APPLICANT: Alsbrook II, John P
TITLE OF INVENTION: No. US20030083244A1e1 Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 15966-783
CURRENT APPLICATION NUMBER: US/09/842,758
CURRENT FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: 60/200,158
PRIOR FILING DATE: 2000-04-26
PRIOR APPLICATION NUMBER: 60/200,613
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,780
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/201,006
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,007
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,236
PRIOR FILING DATE: 2000-05-01

/	PRIOR APPLICATION NUMBER:	60/201,238
/	PRIOR FILING DATE:	2000-05-01
/	PRIOR APPLICATION NUMBER:	60/201,186
/	PRIOR FILING DATE:	2000-05-02
/	PRIOR APPLICATION NUMBER:	60/201,474
/	PRIOR FILING DATE:	2000-05-03
/	PRIOR APPLICATION NUMBER:	60/201,508
/	PRIOR FILING DATE:	2000-05-03
/	PRIOR APPLICATION NUMBER:	60/220,591
/	PRIOR FILING DATE:	2000-07-25
/	PRIOR APPLICATION NUMBER:	60/232,678
/	PRIOR FILING DATE:	2000-09-15
/	PRIOR APPLICATION NUMBER:	60/263,217
/	PRIOR FILING DATE:	2001-01-22
/	PRIOR APPLICATION NUMBER:	60/265,160
/	PRIOR FILING DATE:	2001-01-30
/	NUMBER OF SEQ ID NOS:	113
/	SOFTWARE:	Patentin Ver. 2.1
/	SEQ ID NO 42	:
/	LENGTH:	1661
/	TYPE:	PRT
/	ORGANISM:	Sus scrofa
/	US-09-842-758-42	:
Query Match		
Best Local Similarity 78.1%; Score 6724; DB 10; Length 1661;		
Matches 1272; Conservative 189; Mismatches 200; Indels 2; Gaps 2;		
QY	1	MGPTSGPSLLLLLLLTHPLALGSPMYSIIPNLRLESEETWVLEBAHQGDVPVTVH 60
DB	1	MGSTSGPRLLLLLTSJFALGDPITYIITPNVRLSEEMVLEAHEQGQDIRVSIVTH 60
QY	61	DPPGKKLVLSSEKTVLPATNMGNVTFTIPANKREFSKSGKNKFVTCATFGTQVEKV 120
DB	61	DPAKQVQLSSETTLNNANNYLSTVNKI PASKEFKSKG-KHKFTVQALPGNVQVEKV 119
QY	121	VLVLSQSGLYFIQDKTYITPGSTVLRYIPTNVNHKLTPVRTVMNINENPISGIPVKQDSL 180
DB	120	VLVLSQSGLYFIQDKTYITPGSTVLRYIFTVDHKLTPVGQTIIVTETPEGIDIKRDSL 179
QY	181	SSQNQLGVPLSDWIPELVNNGWKIRAYYENSPOOVFSTEPEVKEYVLPSEVIIVERTE 240
DB	180	SSHQFGILASWNIPELVNNGWKIRAHVEDAPQOVFAEPEVKEYVLPSEVQVEPSE 239
QY	241	KFYIYNKSGLEVITIFARFLYKKVBGTAFVIFGIQDGEQRISLPESUKRPIEDGSQEV 300
DB	240	KFYIIDDPNGVTNVIIFARFLYGESVDATAFIEGVQDQGRISLSQSUTRPVIDGTGEA 299
QY	301	VLSRKVLDDGQNPRAEDLVGSLYSATVILHSGSDMWCAERSGIGPIVTSPYQHFTKT 360
DB	300	TLSQGVLLGHVYSSVNDLVGSKIYVSVTVILNSGSDMWCAERTGPIVTSPYQHFTKT 359
QY	361	KPKFKGPFPDLVFVTPNPDGSPAIRVPVAQGEDTVQSLTQGDVAKLSINTHPSQKPL 420
DB	360	KPKFKAMPFDLMYYVTPNPDGSPARHIPVVYTE-DFKVRSLTDGEGVAKLSINTPNRNL 418
QY	421	SITVTRTKQEELSEAEQATRWQALPYSTVCGSNXNYLHLSVLRTLRPCGETLNVLNRLMD 480
DB	419	PITVTRKDGIPAARQAASKMTWVLPYNTQNSXNYLHLSLPRVELKPGENLVNPHLETD 478
QY	481	RAHEAKTRYTYLINMKGRLLKAGRVREPGQDLVVLPLSITTTDFIPFRLVAYTIIIGA 540
DB	479	PGYQDKIRYTYLINMKGRLLKAGRVREGSQVWWVPLTITTTDFIPFRLVAYTIIIAA 538
QY	541	SGQREVADSVVWDVKDSCVGSLVKSQGESDRQPVPGGQMTLKITEGHDGARVLVAVDK 600
DB	539	NGQREVADSVVWDVKDCSCVTLVKGVGQDQKHRRPGQMTLEIQGERGARVLVAVDK 598
QY	601	GVEVLNKQNKLTOSKINDWVEKADICTPGSGKDYAGVFSDAQLTFTSSSGQOIAQRAEL 660
DB	599	GVEVLNKQNKLTQBRIWDWVEKADICTPGSGKDFAGVFTDAGLAFKSKSGLOTFQRAEL 658
QY	661	OCPQFAARRRSVOLTEKRMDKVGKYPKELRKCCEDCMRENPMRFSCORRTFRFISLGAC 720

APPLICANT: Vernet, Corine A.M.
 APPLICANT: Fernandes, Elma R.
 APPLICANT: Gerlach, Valerie
 APPLICANT: Malyankar, Uriel M.
 APPLICANT: Bolgog, Ferenc L.
 APPLICANT: Zerhusen, Bryan D.
 APPLICANT: Szytek, Kimberly A.
 APPLICANT: Meijnder, Kumud
 APPLICANT: Tchernev, Velizar T.
 APPLICANT: Patturajan, Meera
 APPLICANT: Burgess, Catherine E.
 APPLICANT: Gangolli, Baha A.
 APPLICANT: Smithson, Glenda
 APPLICANT: Rastelli, Luca
 APPLICANT: MacDougall, John R.
 APPLICANT: Taupier, Raymond J.
 APPLICANT: Grosse, William M.
 APPLICANT: Szekeres, Edward S.
 APPLICANT: Alsobrook, John P.
 APPLICANT: Anderson, David W.
 APPLICANT: Guo, Xiaojia (Sasha)
 APPLICANT: Li, Li
 APPLICANT: Zhong, Mei
 TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
 FILE REFERENCE: 15965-783 CIP1
 CURRENT APPLICATION NUMBER: US/10/174,333
 CURRENT FILING DATE: 2002-06-18
 PRIOR APPLICATION NUMBER: 60/193,664
 PRIOR FILING DATE: 2000-03-31
 PRIOR APPLICATION NUMBER: 60/194,614
 PRIOR FILING DATE: 2000-04-05
 PRIOR APPLICATION NUMBER: 60/195,063
 PRIOR FILING DATE: 2000-04-06
 PRIOR APPLICATION NUMBER: 60/195,066
 PRIOR FILING DATE: 2000-04-06
 PRIOR APPLICATION NUMBER: 60/195,067
 PRIOR FILING DATE: 2000-04-06
 PRIOR APPLICATION NUMBER: 60/195,068
 PRIOR FILING DATE: 2000-04-06
 PRIOR APPLICATION NUMBER: 60/195,069
 PRIOR FILING DATE: 2000-04-06
 PRIOR APPLICATION NUMBER: 60/195,070
 PRIOR FILING DATE: 2000-04-06
 PRIOR APPLICATION NUMBER: 60/195,510
 PRIOR FILING DATE: 2000-04-06
 PRIOR APPLICATION NUMBER: 60/219,855
 PRIOR FILING DATE: 2000-07-21
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 186
 SOFTWARE: CuroSequid version 0.1
 SEQ ID NO 42
 LENGTH: 1661
 TYPE: PR1
 ORGANISM: Sus scrofa
 US-10-174-333-42

Query Match 78.1%; Score 6724; DB 15; Length 1661;
 Best Local Similarity 76.3%; Pred. No. 0;
 Matches 1272; Conservative 189; Mismatches 200; Indels 2; Gaps 2;

QY 1 MGSTGSPSLLLTLPLALGSPMYSIITPNILRLESETVWLEAHDAQDGVPTVTVH 60
 DB 1 MGSTGSPRLLLLSLPLALGPIVITIPNVRLSEEMVWLEAHFGQDIRSVTVH 60
 QY 61 DFPQKLVLSSEKTLTPATNMGVTFIPANRFEKSEKGNKXFTVQATFGQVEKV 120
 DB 61 DFPKQVLSSESTTLNANNYSITVNIKIPAKSEKSEK-HREVTQALFGNVQVEK 119
 QY 121 VLVSQSGYFIQTDKTIYTPGSTVLYRIFTVNHKLLPVGRVTVMVNIENPSGIPVKQDSL 180
 DB 120 VLVSQSGYFIQTDKTIYTPGSTVLYRIFTVDHLLPVQGIIVTITPTSGIDIKRDSL 179

181 SSQNLGVLPLSWDIPELVNMGMKTRAYYENSQQVFSTEFVYKZVLPSPSEVIVPTE 240
 DB SSSNQFGLALSWMIPELVNMGMKIRAHYEDAPQVFSAEFVYKZVLPSPSEVIVPTE 239
 QY 241 KFYIYNEKGLVITITARFLYGVKVEGTAFVIGIQDGEORISLPSLKRIPIDEGSGEV 300
 DB KFYIIDDENGTLVNIARFLYGESVDGTAFVIGVGDGORISLSQSLRFLIDGTGEA 299
 QY 301 VLSRKVLDDGVNPRAEEDLVKSLYVSATVILHSGSDMVQABSGPIVTSYQIHFYKT 360
 DB TLSQGVLLNGVHYSSVNDLVKSLYVSATVILHSGSDMVQABSGPIVTSYQIHFYKT 359
 QY 361 PKYFKPGMPFDMVFTNPDGSPAYRVPVAVQGEDTVQSLTQDGVAKLSINTHSQKPL 420
 DB PKFKFPAFDMVFTNPDGSPARHIPPVTE-DFKVRSLTQDGVAKLSINTPDNRSL 418
 QY 421 SITVTKKQELSEABQATRTMQALPYSTVGNNSNYLHSLVRLTELPGTIANVLLRMD 480
 DB FITVTRTXDGIIPAARQASKTMHVLVYNTQNSKNYHLSPRVLEKPGENLVNPHLRTD 478
 QY 481 RAHEAKIRYTYLIMNKGRLLKAGQVRPEQDLVLPISITITDFIPSPRLVAYTYLGA 540
 DB PGYQDKIRYTYLIMNKGRLLKAGQVRPEQDLVLPISITITDFIPSPRLVAYTYLGA 538
 QY 541 SCOREVWADSVWVVKSCVGSVWVKSQSEDQVPGQOMTLKIEGDHAGARVLVAUDK 600
 DB NGQREVWADSVWVVKSCVGTLLVKGKGQKQHRPGQOMTLKIEGDHAGARVLVAUDK 598
 QY 601 GVFLNKKNKLTQSKIWDVVEKADIGCTPGSGDKDYAGVFSADAGLTSSSGQQTQARBL 660
 DB GVFLNKKNKLTORRINDVVEKADIGCTPGSGDKDFAGVFTDAGLAFKSKGQTPQADL 658
 QY 661 QCPQPAARRRSVQTEKMDKVGKPKELKRCCEGMDRENPKRSCQRTAFISLGEAC 720
 DB ECPKPAARRRSVQLEKMDKVGKPKELKRCCEGMDRENPKRSCQRTAFISLGEAC 718
 QY 721 KKVFLDCCNITELRRGHARASHGLARSLNLDIIEENIVSRSEFPESLWNVEDLKE 780
 DB VKAFDCCYEYAKLRQHSRNKPLGLARSLDDEIPEEDIIISRSQFPESLWNTIEFKE 778
 QY 781 PPKGISTKLMNIPLKDSITTWELAVMSDKGICVADPFVTVVQDFFIDRLRPYSV 840
 DB PDKNGISTKLMNVLKDSITTWELAVMSDKGICVADPFVTVVQDFFIDRLRPYSV 838
 QY 841 RNEQVEIRAVLYNRQOELKVRVELLHNPACSLATTTKRHOQTITIPKSSLSVPYVI 900
 DB RNEQVEIRAILYNYREAEKLVVVELLYNPACSLATAKKGHOQTITVPAKSSVPYVI 898
 QY 901 VPLKTGLQEVKAAVHHFISDQVRKSLKVVEGIMKNTVAVRILDPERLQREGVQKE 960
 DB VELKTGLQEVKAAVHHFISDQVRKSLKVVEGIMKNTVAVRILDPERLQREGVQKE 958
 QY 961 DIPPADLSDQVPTSETRILLOQTPVAQMTEDAVDAERLKLHIVTPSGCGEQNMIGMP 1020
 DB EIPPADLSDQVPTSETRILLOQTPVAQMTEDAVDAERLKLHIVTPSGCGEQNMIGMP 1018
 QY 1021 TVIAVHYLDTEQWEKLEKQCALBLIKKGYTOQLAFQPSAAFAAFAKAPSTWLA 1080
 DB TVIAVHYLDTEQWEKLEKQCALBLIKKGYTOQLAFQPSAAFAAFAKAPSTWLA 1078
 QY 1081 VYVVKFSLAVNLIAIDSVLCGAVKWLILEKQKPDGVQFQEDAFVHQEMIGGLNNNEKD 1140
 DB VYVVKFSAVNLIAIDSVLCGAVKWLILEKQKPDGVQFQEDAFVHQEMIGGLNNNEKD 1138
 QY 1141 MALTAFLVLSIOBAKDICEQVNSLPGSITKAGDFLEANYMLQSVTVVAIAGYALQMG 1200
 DB VSLTAFLVLSIOBAKDICEQVNSLPGSITKAGDFLEANYMLQSVTVVAIAGYALQMG 1198
 QY 1201 ELKGPLLNKLTITAKDNREDPGQKLYNVEATSYALLALLQLKDFDVPVVRVWLNOR 1260
 DB ELKGPLLNKLTITAKDNREDPGQKLYNVEATSYALLALLQLKDFDVPVVRVWLNOR 1258
 QY 1261 YVGGGSGTQTFWVQALAQYQKADPHOELNLDVSLQPSRSSKITHIHWEASALLR 1320

Db 1259 YGGSYGTQATFWFOALAQYQKQVDPDHKDLNLDVSIHLPSESAPVRHRIILKESALLR 1318
Qy 1321 SEETKEGEGTVAAGGQOTLSVWYHAKAQOJTCNKFDLKVTKPAPETERPQDA 1380
Db 1319 SEETKEGEGTVAAGGQOTLSVWYHAKAQOJTCNKFDLKVTKPAPETERPQDA 1378
Qy 1381 KNTMILICIRYRGDQATMSIILDISMTGFAPDQDOLQALANGVDVYISKYELDKAFSD 1440
Db 1379 KSSMVLICIRYRGDQATMSIILDISMTGFAPDQDOLQALANGVDVYISKYELDKAFSD 1438
Qy 1441 KNTMILICIRYRGDQATMSIILDISMTGFAPDQDOLQALANGVDVYISKYELDKAFSD 1500
Db 1439 KNTMILICIRYRGDQATMSIILDISMTGFAPDQDOLQALANGVDVYISKYELDKAFSD 1498
Qy 1501 KNTMILICIRYRGDQATMSIILDISMTGFAPDQDOLQALANGVDVYISKYELDKAFSD 1560
Db 1499 KNTMILICIRYRGDQATMSIILDISMTGFAPDQDOLQALANGVDVYISKYELDKAFSD 1558
Qy 1561 KNTMILICIRYRGDQATMSIILDISMTGFAPDQDOLQALANGVDVYISKYELDKAFSD 1620
Db 1559 KNTMILICIRYRGDQATMSIILDISMTGFAPDQDOLQALANGVDVYISKYELDKAFSD 1618
Qy 1621 KNTMILICIRYRGDQATMSIILDISMTGFAPDQDOLQALANGVDVYISKYELDKAFSD 1663
Db 1619 KNTMILICIRYRGDQATMSIILDISMTGFAPDQDOLQALANGVDVYISKYELDKAFSD 1661

RESULT 9

US-09-925-442-2
; Sequence 2, Application US/09925442
; Patent No. US20020103346A1
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; BREDEHORST, REINHORST
; KOCK, MICHAEL
; FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVF
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/925,442
; FILING DATE: 10-Aug-2001
; CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/017,947
; FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-0107-0X

TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220

INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1642 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-925-442-2

Query Match 49.1%; Score 4224; DB 9; Length 1642;
Best Local Similarity 49.6%; Pred. No. 0;
Matches 824; Conservative 324; Mismatches 478; Indels 36; Gaps 20;

Qy 9 LLLLLLTHPLALGSPMYSIIITPILRLSEETMVLAEHAQDQDPVVTVHDFPKKLIV 68
Db 8 LVAALLIGPSSGSHGALYTLITPAVLRTDTTEQILVEAHGDSPTKQLDIFVHDFPKKQT 67
Qy 69 LSEKTVLTPATNMGNVTFTIPANREFKSEKGRKNFVTVQATFTQTV-VKXVLVLSQS 127
Db 68 LFTQVDMNPAGGMLVTPTEIPA-KEVSTDSRONQYVVVQVT-GPQVRLEKWLSSQS 125
Qy 128 GYLFPIQDTKIYTPGSGTVLYRIPTVNNKLLFPVGRVTMWNENPEGIPVKQDLSLSONQLG 187
Db 126 SFLPIQDTKGIYTPGSPVLYRVFSDHNTSKNKTIVIVEFQTPGEGILVSSNSVD----- 179
Qy 188 VLPLSW--DIPELVNGQWKIRAYTENSQQVSTFEFEVEYVLPSPFEVTEPKFYI 245
Db 180 -LNFPPYNNLPDLVSLGTWRIIVAKYHSPEN-YTAYDPVKYVLPSPFEVRLQPESEKFFYI 237
Qy 246 YNEKGLVITITAFRLYKKEVEGTAFVIFGIQDEQRISLPESLKRIPETEDSGSEVILSRK 305
Db 238 DGNENFHVISITARYLYGEEVEGVAFVFGVKIDDAKKSIPDSLTIRIPIIDGDKATLKD 297
Qy 306 VLLDGVQNPRAEDLVGKSLYVSATVILHSGSDMVAERSGPIVITSPQIHFHTKPKYK 365
Db 298 TFRS--RFFNMLNELVHTLYASVTYNTESGSDMVAETESGHIHIVASPYQIHFHTKPKYK 355
Qy 366 PGMPFDLMVFTNPDGSPAYRVPVAVOGEDTVQSITQDGVAKLSINTHPSKPLSITVR 425
Db 356 PGMPYELTVVYVNPDPGSPAHPVYSEAFISM-GITLSDGAKLINIPLNAQSLPIIVR 414
Qy 426 TKQELSEABQATRTMQLPYSTVGNSSNNYHLVRLTELPGETLNVNLLMRDRAHEA 485
Db 415 TNHGDLPREQATKSMITAYQTGGSGNYLHVATITSEIKPDNLNPNVKNANSLK 474
Qy 486 KIRYTYLIMNKRILKAGRVREPGQDLVPLSTITDTPSPFLVAYVYTLIGASGORE 545
Db 475 QIKFYIILINKGIFKVGQPRRDQONLVNHLHTPDLIPSPFRVAYQV----GNNE 530
Qy 546 VVADSVWVDKSCVSLVYKSGQSDROPVPGQONTLKISGDHGRVVLVAVDKGVFL 605
Db 531 IVADSVWVDVKOTCMGTLYVK--GDNLIQMPGAAMKIKLEGDGPARVGLVAVDKAVYVL 587
Qy 606 NKKNKLTQSKINDVVEKADIGCTPGSGDYAGVFDAGLTFTSSSGQQTAKRAELQCPQ 665
Db 588 NDYKISQAKINDTIEKSDFGCTAGSGQNNLGVFEDAGLALTSTNLNTKQESAACKQP 647
Qy 666 AARRRR-SVOLTEKRVKVKY-KPELRKCCEDGRENPMRPFSCORRTFTSLGBACKKV 723
Db 648 ANRRSSVLLLDNSAKAAEFQDQDLKCCEDVNHENPMGYTCCKRAKYQEGDACKAA 707
Qy 724 FLDCCNVITELRQHARASHLGLARSLDELIIAENIVSRSEFPESMLWVDEL-KEPP 782
Db 708 FLECCRYIKGVRENQRESELEFLARDNEDGFIADSDIISRSDFPKSNLMTKDLTEBN 767
Qy 783 KNGISTKLMMIFLKDSITTWELAVMSDKKICVADPFVTVMDFFIDRLPYSVVRN 842
Db 768 SGGISSTKMSFYLDSITTWVAVLSFTPTKICVAEPYERVMKVFDFIDQMPYSVVRN 827
Qy 843 EQVEIRAVLYNRQOELKVRVELLHNPACFSLATTKERHQOTITIPKSSLSVYVIVP 902
Db 828 EQVEIRAILHNY-VNEDIYVRVELLYNPAFCASATKQRYQQQFIIKALSSRAVPFVIVP 886
Qy 903 LKTGLQEVVEKAAVYHFIISDGVKSLKVVPPGIRMNKTAVRITLDPRLREGVQKEDI 962
Db 887 LEQGLHDEIKASVQBALMSDGVKRLKVVPEGVQ-KSIVITVKLDPRAKGVGGTQLEVI 945
Qy 963 PPADLSQVDPDTESETRILLQGTTPVAQMTEDAVDAERLKLHIVTPSGGCEQNMIGMTPTV 1022

Db 946 KARKLDRVDPTEIEKIIIOGDPVQAIIENSIDGSKLNLHIIITPSCGGEQNMIRMAAPV 1005
Qy 1023 IAVHYLDETEQMEKFKLEKROGALELIKGYTQQLAFQPSAFAAFVKRAPSTWLTAYV 1082
Db 1006 IATYVLDTEQNETLGINRTEAVNOIVTVGAQVYKKAADSHSYAAFTNRASSSMLTAYV 1065
Qy 1083 VKYFSLAVNLIA-IDSOVLGCAVKWILILEKQPDGDFQEDAPVHOEMIGSLRNNKDM 1141
Db 1066 VKYFMAAKVAGISHEIICGGVRWLLNRQPDGAFKENAPVSGTQGGIQ-GABEEV 1124
Qy 1142 ALTAFLVLSIQEAKDICEBQVNSLPGSITKAGDFLEANYMLORSYVAIAGYALAOQMR 1201
Db 1125 YLTAFLVALLESKTIENDVNSLDSISKATNYLLKKYKLOQPYTTALTAYALAAQ 1184
Qy 1202 LKPLINKELTTAKDNKRWEDGKQLYNVEATSYALLALLOLKDPDFVPVVRWLNQRY 1261
Db 1185 LNDRLVMAASTGRD--HWEENYNAHTNIEGTSYALLALLKWKPKDQGTGPIVRWLTQNF 1242
Qy 1262 YGGYGSTQATFVFOALAQYKQDAPDHOELNLDVSLQPSRSSKITHRIHWSASLRS 1321
Db 1243 YGTYGOTCATVVAFOALAEYELQMPHDKDNLDTIELPDREVPYRINYENALLART 1302
Qy 1322 EEPKENEGETVTAEGKQGTLSVWYHAKADQOL-TCNKEDLKVITKIPAPETEKRPQDA 1380
Db 1303 VETKLNQDITVATAGGKATWILTIFYNAQLQEKANVCNKFHLNVSE---NIHLNANGA 1359
Qy 1381 KNTMLEICTRYRGDODATMSILDISMGTGAPDITDLDKQLANGVDRIYSKYELDKAFSD 1440
Db 1360 KGALMLKICTRYLGEVDSWTIIDI:SMLTGFLPDABEDLTRLSKGVDRYISRYEVDNMAQ 1419
Qy 1441 RNTLLIYLDKVSHESDCLAFKHQYFENVLLTOPGAVKYAYVYMLESCTEFYHPEKEDG 1500
Db 1420 KVAVIYLNKVSHEDECHLKHFEVGFQPGSVKYSYNDLDEKTFYHPDGTG 1479
Qy 1501 KLNLKRCDELCAEBCNFIQKSDDKVTLEELERKACBPYDVYVYKTLKVKQLSNDFDE 1560
Db 1480 LLNKICIGNVCRAGETCSLNLHQRIDVPLQIEKACETNVYVYKTLKRIEQDGNDI 1539
Qy 1561 YIVAIQTKSGDEVOVQQRFTFSPIKREALKLEKHYLMWGLSSDFWGBKPNLSY 1620
Db 1540 YVMDVLEVIKQGTDENRARTHQYISQRKQCALNLKVNDDYVLINGSRDLPTDKLISY 1599
Qy 1621 IIGKDTWHERWPEDESCQENKQKQODLGAFTESMVYFGCP 1662
Db 1600 IITKNTWIERWPEDESCQEEFQKLCDDPAQPSYLTTEFGCP 1641

RESULT 10

US-09-925-442-35
; Sequence 35, Application US/0925442
; Patent No. US20020103346A1
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; BREDEHORST, REINHORST
; KOCK, MICHAEL
; FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVF
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUBSTADT,
; P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/925,442

FILING DATE: 10-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/017,947
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORVAN P.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-0107-0X
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 1648 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-925-442-35

Query Match

Best Local Similarity 48.7%; Score 4194; DB 9; Length 1648;
Matches 821; Conservative 324; Mismatches 481; Indels 36; Gaps 20;

Qy 9 LLLALLLHPLALGSPMYSIITNLRLESEETMVLFAHDAQDGPVTVTHDFPKKLV 68
Db 8 LVAALLIGFPSSHGALYTLITPAVLTDTDEQILVEAHGDSPTKQLDIFVHDFPRKQKT 67
Qy 69 LSSEXTVLTATNMGVNTFTIPANREFKSEKGRNFVTVCATGCTQV-VEKVLVLSQS 127
Db 68 LFQKRVNDKXNAGGMLVTFTEIPA-KEVSDTSQNVVVQVT-GQVLEKVLVLSQS 125
Qy 128 GYLFTQDKTITPGSTVLYRIFTVNHKLPLVGTVMNIENPEGTPVKQDLSNQNLG 187
Db 126 SFLFTQDKGYTTPGSPVLYRVFSMDHTSKMKTIVBEPQTEGLVLSNSVD----- 179
Qy 188 VLPLSN--DIEPLNMGWKIRAVYNSPOOVSTFEVKEYVLPSEFVIVBTEKFFVY 245
Db 180 -INFPWYNLDELVLSTGTWRAVAKYHSPEN-VTAIFDVRYVLPSEFVRLQPSKFFVY 237
Qy 246 YNEKGLVITITARELYGKKVEGTAFVIFGIQDGEQRISLPESLKRPIEDGSGEVVLSRK 305
Db 238 DGNENFHVSIITARYLYGEEVGAFLVFGVKIDAKKSIDSLTRIPIDGDKATLKED 297
Qy 306 VLLDGVQNPRAEDLVGSLYSUSATVILHSGSDMVQAEKSGIPVTSPIYHFTKPKYFK 365
Db 298 TFRS--RFPNLELVGHTLYASVTVMTESGDMVVTESQGIHVIVASPYQHFTKPKYFK 355
Qy 366 PQMPFDLMVFTNPDGSPAYRVPVAVQGEDTVQSLTQGDGVAKLSINTHPSQKPLSITVR 425
Db 356 PGMPYELTVVYTNPDGSPAAHVPPVSEAFHSM-GTLLSDGTAKLILNPLNAQSLPITVR 414
Qy 426 TKKQELSEAEQATMCAIPYSTVGNNSNYLHLVSLTELPCGTNNVFLMRDRAHEA 485
Db 415 THGDLPRERATKSMATAIAQTGGSGNVLHVVAITSEIKPGDNLFPVNFVKNANGSLK 474
Qy 486 KIRYVYILMNKRLKAGQVREFGQDLVVLPLSITDTPFSRFLVAVYVTLIGASQRE 545
Db 475 QIKYFTYILNLKGIKFKVGRQPRRDGQNLVTMNLHITPDLPISFRFVAYYQV---GNNE 530
Qy 546 VVADSVWVDVKDSCVGLSVKSGSEDRQVPVQOQMTLKIEGDHGVAVLVAVDKGVFL 605
Db 531 IVADSVWVDVDTCTNGTLVVK--GDNLIOMPAAMKIKLEGDFGARVGLVAVDKAVYL 587
Qy 606 NKENKLTOSKIWDVVEKADIGCTFGSGKDVAGVPSDAGLTFTTSSSQOQTAQRAELQCP 665
Db 588 NDKYKISQAKIWDITIEKSDFGCTAGSQNNLGVFEAGLALTSTNLNTRKQSAKCPQ 647
Qy 666 RARRR-SVQJTEKRMKDKVGY-PKELURKCCEDGMENPMRFSCORRTFISIGEACKV 723
Db 648 ANRRRSVLLDSDNASKAEAFQDQDLRKCCEDVMHNPNGYTCERAKYIQEGDACKAA 707

```

Qy 724 FLDCNVTETRRCHASHGLARSNLDDEIIAENIUSRSSEPPSSMLNVEDL-KSP 782
Db 708 FLECRKIKGRDNRESELSFLARDNEDGFTADSDIISRSDFPKWMLUKTDBEPN 767
Qy 783 KNGISTKMLNIFLKDSTTTWEIIAVMSDKKG:CVADPFVETVMQFFIDRLPLSYVRN 842
Db 768 SQGISSTMGPFYLRDSTTTWVLAVSFTPKGICVAEPYEURVMKVFIDLQMPYSVK 827
Qy 843 EQVEIRAVLYNROQELKVRVELLHNPFCSLATTKRRHQOITIPPKSSLSVPVIVP 902
Db 828 EQVEIRAILHNY-VNEDIYVRVELLYNPFCSTAGKQYRQOPPIKALSSRAVFFVVP 886
Qy 903 LKTSQVEVEKAAVYHHFISGVKRSUKVYVEGIRMKNTVAVRTLDPERLREGVCREDI 962
Db 887 LEQGLHVDVEIKASVQEAALSDGVRKLLKVVPEGVQ-KSIVTIKLDPRAKGVGGTQLEVI 945
Qy 963 PPADLSQVPTSEETRIILQGTFAQMTEDAVDAERLHLIVTPSGCGEQNMIGWTPV 1022
Db 946 KARKLDRVPDTEIETKIIIQGDPVAQIENSIDGSKLNLHIIITPSGCGEQNMIRNAPV 1005
Qy 1023 IAVHYLBETEOWEKFLEKQGALELIKKGYTQCLAPROPSSAFAPVKRAPSTWLTAYV 1082
Db 1006 IATYLDTTTEQWETLGINRRTEAVNQIVTGYAQMVKKADSHSYAAFTNRASSWLTAYV 1065
Qy 1083 KVPFSLAVNLIA-IDSOVLGAGVKWLLILEKQKPDGVFOEDAPVHOEMIGGLRNNEKDM 1141
Db 1066 KVPFAAKWVAGISHIIICGVWVLLINRQQDPGAFKNAPVLSGTMOGGIQ-GABEEV 1124
Qy 1142 ALTFVLISQBAKIDICEEOVNSLPGSITKAGDFLEANYMNGORSYTVIAIAGYALQMR 1201
Db 1125 YLTFILVALLIESKTCNDVYNSLDSIKKATNLLKYEKLQRPYTTALTAYALAAADQ 1184
Qy 1202 LKGPLLNKFLITAKDKWRDPPGKQLVNVEATSYALLALLQKDPDPVPPVVRWLEORY 1261
Db 1185 LNDRLVMAASTGRD--HNEYVAHNIESTSYALLALLKMKFFQOTGPIVRWTDQNF 1242
Qy 1262 YGGYSGTQATFMVFOALAOYQKADAPHOELNDVLSQLPSRSSKITHIHWSASLRS 1321
Db 1243 YGETYGTQATWAFQALAEYEQMPTHKDLNDJITIELPDRVPIRYINYNALLART 1302
Qy 1322 EETKENEGFTVTSAGKGGTLTVVVMYHAKADOL-TCKNFDLKVTKZAPETERKPODA 1380
Db 1303 VETKLNQDIIVTASGDGKATWLTITFNAOLQSKANVCKHNLNVSVE---NIHLNAMA 1359
Qy 1381 KNTMLEICTRYRGDQATWSIIDISMGTGAPDPTDQLQANGVDRIYSKYELDKAFSD 1440
Db 1360 KGALMLKICTRYLGEVDSTWITIDISMLTGFLPDAEDLRLSGVDRIYSRYEVDNNMAQ 1419
Qy 1441 RNTLIIYLDKVSHEDEDLAFKVOHYENVNELIQGAVKVVAYYNNLESCSTRYHPKEDG 1500
Db 1420 KVAVIIYLNKVSHEDEDLAFKHPKILKHPVGFQIPGSVKVSYNNLDEKCTKYHPDKGTG 1479
Qy 1501 KLNKLCRDELRCRABENCFIQSDDKVTLERLDKACEGVDVYKTRLVKQLSNDPDE 1560
Db 1480 LLNKICIGNVCRCAGETCSSLNQERIDVPLQIEKACETNVDVYVYTKLRIEEQDGN 1539
Qy 1561 YIMAEOTIKSGSDEVQVQORTFISPKCREALKLEBKHYLMWGLSDEWGEKPNLSY 1620
Db 1540 YVMDVLEIKGTDENPRAKHOVIQSKQCEALNKVNDVYLINGSRSDLLPTKDKISY 1599
Qy 1621 IIGKDTVVEHWPBEDEQDENOKQODLGAFTESMVVFGCP 1662
Db 1600 IITKNTIWRPHEDEQCEBEFQKLDFFAQFSYTLTFEGCP 1641

```

RESULT 11

```

US-09-925-442-34
; Sequence 34, Application US/09925442
; Patent No. US20020103346A1
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; BREDEHORST, REINHORST

```

```

; KOCK, MICHAEL
; FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVF
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/925,442
; FILING DATE: 10-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/017,947
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-0107-OX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1333 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-09-925-442-34

```

Query Match 41.1%; Score 3536; DB 9; Length 1333;

Best Local Similarity 51.2%; Pred. No. 9.2e-272;
Matches 683; Conservative 254; Mismatches 372; Indels 24; Gaps 15;

```

Qy 334 SGSDMVOAERSGIIPIVTSYQ:HFHTKTPKYPKPGMPFDMVMVTNPDGSPAYRPV--A 390
Db 3 SGSDMVMTEQSGIHIIVTSYQIYFTKTPKYPKPGMPFELTVVYVTRKPDGSPAHPVWSEA 62
Qy 391 VQGEDTVQSILQTQGDGVAKLSINTHPSQKPLSTTVATKQELSEAEQATRTMOALPYSTVG 450
Db 63 IHSEGT-----TISDGTAKLFINTPQAGSLPITVTNHDLPREQAISKMTATAYQIQ 118
Qy 451 NSNNYLHSLVRLTELRPGSETLVNLFLLMRDRAHEAKIRYYTYLIMNKRLKAGQVREP 510
Db 119 GSGNYLHVAITSTEIKPGDNLVFNFNVRGNANSLAQIKYFTYLLINKGKIFKVRQHRGD 178
Qy 511 GDLVLPISITTDTPSPFRLVAYTYTLIGASQSRV-VADSVVDVKDSVCGSLVVKSGQ 569
Db 179 G-NLVTNWLHIPTLIPSPFRVAYTVQV----GNNEIEVADSVVDVKDTCGTLYVKGAT 233
Qy 570 SED-RQFVPGQQTMLKIEGDHGAVVVLVAVDKGVFLNKKKLTQSKLWVVEKADIGCT 628
Db 234 SRDNRIQMPGAAMKIKLEGDPGAWIGLVAVDKAEVYLDNDKYKISQAKTWDTIKSDFGCT 293
Qy 629 PGSGKDYAGVFS DAGLTFTSSSGQQTQAPAEQCPQPARRRR-SVOLTERRMDKVGKY- 686
Db 294 AGSGQNNLGVFEDAGLALTSTNLTNKQSAKCEQPARRRRRSVLLDSDNASKAQFQ 353
Qy 687 PKELRYCCEDGMRENPMRFSQQRTRFISLGBACKKVKVLDCCNYITELRRQHAASHLGL 746
Db 354 DQDLRKCCEGTGMHENPMGHTCEKREKVIQEGDACKAAFLKCHYIKGIQDDNRESEFL 413

```


QY 1594 FISPIKREALKLEBKHYLMWGLSDFWGEKENLSYIIIGKOTWVEHWPEDECCDEBNQ 1643
DB 626 FISPIKREALKLEBKHYLMWGLSDFWGEKENLSYIIIGKOTWVEHWPEDECCDEBNQ 685
QY 1644 KOCQDLGAFSTSMVVFQCPN 1663
DB 686 KOCQDLGAFSTSMVVFQCPN 705

RESULT 13

US-09-842-758-6
; Sequence 6, Application US/09842758
; Publication No. US20030083244A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Spytke, Kimberly A.
; APPLICANT: Majumder, Kumud
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gangoli, Esha A.
; APPLICANT: Smithson, Glennda
; APPLICANT: Rastelli, Luca
; APPLICANT: Macdougall, John R.
; APPLICANT: Taupier, Raymond J.
; APPLICANT: Grosse, William M.
; APPLICANT: Edward, Szekeres S.
; APPLICANT: Alsobrook II, John P.
; TITLE OF INVENTION: No. US20030083244A1: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-783
; CURRENT APPLICATION NUMBER: US/09/842,758
; CURRENT FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/200,158
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/200,613
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,780
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/201,006
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,007
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,236
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,238
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,186
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: 60/201,474
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/201,508
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/220,591
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/232,678
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 60/263,217
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 60/265,160
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 6
; LENGTH: 1612
; TYPE: PRT
; ORGANISM: Homo sapiens

; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1226)
; OTHER INFORMATION: Wherein Xaa is any amino acid.
US-09-842-758-6

Query Match 33.2%; Score 2857.5; DB 10; Length 1612;

Best local similarity 36.8%; Pred. No. 1.6e-217;
Matches 644; Conservative 314; Mismatches 562; Indels 231; Gaps 39;
QY 1 MGPTSGPSLLLLTHLPALGSPMYSIITENILRESEETWVLEAHDAQGD-----VPV 55
DB 1 MSPSLPFSALISLCLSHLPWLPSPRYILVTPVLAVGSPESIHIIQAHSDSQPLRLTKV 60
QY 56 TVTVHDFPGKGLVLSSEKTLVTPATHMGNVTFTIPANREEFKSEKGRNKFVTVQATF--- 112
DB 61 NLTVWDFPMKRTVILARSQILSPGNFMQADVPVPSLMTYLPKPGQ-QYVILRAIWAPT 119
QY 113 -GTQVVEKVLVLSGSGYLFIQTDKIITYTPGSTVLVRIFTVNHKLPLVGRVWVNIENPE 171
DB 120 SGSSFEKMWLVALHAGYFIQTEKTIYTPSPLVHYRVFTVNHKMDPVTRITFLDKNPD 179
QY 172 GIPVKQDSLSSQNLGLVPLSWDIP-----LVNMGMKIRAYVENSPOOVSTEFVEKVV 228
DB 180 GSPASRVLVHSQDQPGVGLASWG-PEKSLCLSGTWTIERASYQSTPKKFAAFDVKVY 238
QY 229 LPSFEVIVPTEKPYIYNEKGLVITITARFLYKKVEGTAPVIFIGIQDGPQRISLPESL 288
DB 239 LPSFEVQLVFNKTFYFLKDE-ALGVDIQARYIFNKFPVDGHALVIFGVKLDSCRPIQSSL 297
QY 289 KRIPIEDSGSEVLSRKVLDDGVQNPRAEDLVKGLVYSATVILHSGSDMWQAERSGPI 348
DB 298 QRVEVTE-----TGEMVQAETSQVKI 319
QY 349 VTSPIQIHFTKTKPKPKPQMPDL-----M 373
DB 320 IQSPYNIKTRTPQVFPKPGMPPHFRVAVVQSSPIQIIFQSHLSHQATAGSFLLPQIPPQ 379
QY 374 VFVTPDGGPAYVPVAVQGEDTVQSLTQGDGVAKLSINTHPSQKPLSITVTRTKQELSE 433
DB 380 VFISNPDGSPASRVLVHSQDQ---KYVTSAGELATLINTDANLDKGLFVKT-EESLQP 435
QY 434 AEOATRTMQALPYSTVGNNSNYLHLSV--LRTELPGETLVNVLNLLRVDRAHEAKIRYTT 491
DB 436 EQASAKMTAWPYLTQDGSNFLHIEVKTGLTEVGSSICLSLN-TRHQDPKTKDKITHT 494
QY 492 YLIAMNKGRLKAGQVRPEPQDLVPLSLITDFTPSRLVAYTL-TGASQQRVWADS 550
DB 495 ILVREGKARQGRQVAVG-----VPSFRLAFYLLPRAEQDDELVADS 540
QY 551 VWVDVKDSCVGLWVKSQSEDR---QPVFGQOMTLKIEGDHGRVVLVAVDKGVFLNKK 607
DB 541 IWIDVNDRCIG--LKVGLKNDRFQSLPNSQVELKVTGDAEATVGLVAVDKAVYVLS 597
QY 608 KXKLTOSTKWDVVEKADIGCTPGSGKVAGVFSAGLFTFTSSGQOTQRAELQCP----- 663
DB 598 KHLTKQKVNWVVEHDIGCTGSGKDRPAVFKDAGLDLKISTGWDNSHQSHSCQEAEV 657
QY 664 ----QPAARRRRSVQLTEKRMKVGVKPKXL-RKCCEDGMRENPMRFSCQRTFISIGE 718
DB 658 GESLEPGROKRLRSAAETAPLH-SSVNNKFTLEBQKCEAGLRESVGLSCERTVHRHGP 716
QY 719 ACKKVFLDCCNYITELRRQHRASHLGLARSLNDEDI AEENIVSRSEFPESLWNVDEL 778
DB 717 ACVAFLDCCSHL-----LPPADEEDFDDLFLDDMPVRTLFPESWLK----- 759
QY 779 KEPPKNGISTKLMNIFLKDSITTWELAVSM-SDKKGICVADPPFVTVNQOFFIDRLRPY 837
DB 760 ----NSISHYPTISVKVPSITTWQVWVYVSLKAGQGLCVSDPPELTVMKSEFFVDLKLPS 814
QY 838 SVVRNEQVEIRAVLYNVRQNELKVRVELLHNPAFCSLATTKRRHQQTITIPKSSLSVP 897
DB 815 SVIRNEQVQIQAOLVYNFRDQ-AKVRVEPFPKHEITLCSAKPGCAPSHQVVPVPTSSKLVH 873

Db 708 KVSHSETECVGFKHQDFEVLGLOPASVKYDYDEQCTAFY 751

Search completed: December 22, 2004, 00:49:58
Job time : 199 secs

```

; PRIOR APPLICATION NUMBER: 60/242,611
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,612
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,880
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,881
; PRIOR FILING DATE: 2000-10-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 80
; LENGTH: 751
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: Alpha-2-macroglobulin family Consensus Sequence
US-09-981-151A-80

Query Match      21.6%; Score 1855.5; DB 10; Length 751;
Best Local Similarity 52.4%; Pred. No. 3.2e-138;
Matches 400; Conservative 105; Mismatches 220; Indels 39; Gaps 10;

Qy 756 IAEENIVSRSEPPESLWVDELKEPPKNGISITKLMNIFLKDSTTTWELLAYSMSDKKI 815
Db 1 IDEDDITIRSYFFESWLWEVEEDRSP-----VLTVNITLPPDSITTWELLAYSMSDKKI 55
Qy 816 CVADPEVITVQDFFIDLPLSYVVRNEQVEIRAVLYNQQLKVRVELLHNP----A 871
Db 56 CVADPEVITVQDFFIDLPLSYVVRNEQVEIRAVLYNQQLKVRVELLHNP----A 871
Qy 872 FCSLATKRRHQTIIPKSSLSVDPVTPKLTGLOVEVKAAYHFI SDGVKSLKV 931
Db 116 FCSLATQRTSSQVAPKSLSSVSVVPLASGLSVEVVASVPEFFVKDAVVKLKV 175
Qy 932 VPESGIRMNKTVAVRTLDPRLGREGVQKEDIPPADLSQVDPDTESETRILLQGTPTVAQMT 991
Db 176 EPESGARKEETVSLLLPPEHLG--GGLVSEVPALKLPDDVPDTEAEAVISVQGDPPVAQAI 234
Qy 992 EDAVDAERLKHLLVTPSGGEGNQMGTTTVAHYLDSTEQWEXFGLKQGALELIK 1051
Db 235 QNTLSGEGNLLRLPSGGEGNQMIYAPTVYHYLDSTEQWEXFGLKQGALELIK 294
Qy 1052 GYTQQLAFROPSSAFVVRAPSTWLTAYVVKVFLAVNLTAIDSVLQCGAVKWLILEK 1111
Db 295 GYQRLNRYKADGSYAAFLHRASSTWLTAFVLKVFQARNYVFI DEEHICGAVKWLILNQ 354
Qy 1112 QKPDGVFOEDAPVTHQEMIGGLNNEKDMALTAFVLISLOEAKDICEQVNSLPGSITK 1171
Db 355 QKDDGVFREGSPVHNEMKGGVGDAAVEVTLTAFITALLLEAKLVCSIPVVALSILK 414
Qy 1172 AGDFLEANYNMLORSYTVAIAGYALAQMG--RLKGPL--LNKFLTTAKDNRMEDPKG- 1225
Db 415 ASDYLLNENYANGQVVTTLTAYALAGVLHKLKEILKSLKEELYKALVKGHWERPQKP 474
Qy 1226 -----QLYVVEATSYALLALLOIKDF---DFVPPVVRWLNERYGGYGSTQ 1270
Db 475 KDAFPHYSPQPOAAAVEMTSYALLALLTLFPKVEAPKVRKWLTEQQYTGFGFGSTQ 534
Qy 1271 ATFWVFOALAQYOKDAPDHOELNLDVSLQLPSSRSKITHRIHWSASLLRSEETKENEFG 1330
Db 535 DTVMALQALSKYGIATPHTHEKNSLVTIQSPSGSFKSHFQILNNAFLRLRPVELPLNEGF 594
Qy 1331 TVTAEGKGCGTIVTWYHAKAKDQLTCKNFKLVTIKPAPETEKPODAKNTMLRICT 1390
Db 595 TVTAKVTGQGLTILVITYRYKVLCKNTFCFDLK--IETVPDICEVPEPKAKNSDYLSICT 652
Qy 1391 RYRGDQ-DATWSILDISMNTGPAFTDDDLKQLANGVDRIYSKLYELDKAFSDRNTLIYLD 1449
Db 653 RYAGSRSDSGMALADISMLTGFTPLKPLKLENGVDRIYSKYEI-----DGNHVLAYLD 707
Qy 1450 KVSHSEDDCLAFKVHQYFNVVELLQPGAVKVAYVYNLEESCTRFY 1493

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2004, 00:29:54 ; Search time 45 Seconds
(without alignments)
2450.819 Million cell updates/sec

Title: US-09-875-519a-22

Perfect score: 8609

Sequence: 1 MGPTSGPSLLLLLLTHPLA.....KQCQDLGAFTESMVFGCPN 1663

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCFUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfileesi.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8609	100.0	1663	2	US-08-793-126-1
2	8609	100.0	1663	3	US-09-132-271-1
3	8609	100.0	1663	3	US-09-142-334-22
4	4389.5	51.0	1651	1	US-08-447-411-2
5	4224	49.1	1642	2	US-08-662-227-2
6	4224	49.1	1642	3	US-09-017-947-2
7	4224	49.1	1642	4	US-09-925-442-2
8	4208	48.9	1642	1	US-08-447-411-45
9	4194	48.7	1648	2	US-08-662-227-35
10	4194	48.7	1648	3	US-09-017-947-35
11	4194	48.7	1648	4	US-09-925-442-35
12	3359	41.3	1333	1	US-08-447-411-76
13	3336	41.1	1333	2	US-08-662-227-34
14	3336	41.1	1333	3	US-09-017-947-34
15	3336	41.1	1333	4	US-09-925-442-34
16	3255	37.8	1643	1	US-07-779-172A-3
17	1840	21.4	1676	3	US-08-487-283A-2
18	1287.5	15.0	308	4	US-09-582-761B-26
19	1287.5	15.0	330	4	US-09-582-761B-37
20	1287.5	15.0	929	4	US-09-582-761B-27
21	1057	12.3	403	1	US-08-118-674-1
22	1016	11.8	1474	4	US-09-241-606-2
23	659	7.7	134	4	US-09-513-999C-4153
24	415	4.8	77	4	US-09-708-606-1
25	407	4.7	76	1	US-08-284-022-1
26	407	4.7	77	3	US-09-246-500B-1
27	339	3.9	63	1	US-08-447-411-24

28	339	3.9	63	1	US-08-447-411-63	Sequence 63, Appl
29	339	3.9	63	2	US-08-662-227-20	Sequence 20, Appl
30	339	3.9	63	3	US-09-017-947-20	Sequence 20, Appl
31	339	3.9	63	4	US-09-925-442-20	Sequence 20, Appl
32	314.5	3.7	78	3	US-09-246-500B-5	Sequence 5, Appl
33	302	3.5	63	1	US-08-447-411-26	Sequence 26, Appl
34	301	3.5	63	1	US-08-447-411-25	Sequence 25, Appl
35	291	3.4	63	1	US-08-447-411-27	Sequence 27, Appl
36	288	3.3	77	3	US-09-246-500B-2	Sequence 2, Appl
37	288	3.3	106	2	US-08-563-186-33	Sequence 33, Appl
38	283	3.3	317	4	US-09-270-767-43623	Sequence 43623, A
39	282.5	3.3	78	3	US-09-246-500B-3	Sequence 3, Appl
40	276.5	3.2	78	3	US-09-246-500B-4	Sequence 4, Appl
41	203	2.4	339	4	US-09-270-767-45607	Sequence 45607, A
42	199	2.3	250	4	US-09-241-606-4	Sequence 4, Appl
43	187	2.2	145	4	US-09-270-767-59001	Sequence 59001, A
44	186	2.2	63	1	US-08-447-411-23	Sequence 23, Appl
45	183.5	2.1	3878	4	US-09-914-259-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-08-793-126-1
; Sequence 1, Application US/08793126
; Patent No. 5843297
; GENERAL INFORMATION:
; APPLICANT: Harrison, Richard Alexander
; APPLICANT: Farries, Charles Timothy
; TITLE OF INVENTION: MODIFIED HUMAN C3 PROTEINS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HALE AND DORR LLP
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: United States of America
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,126
; FILING DATE: 07-FEB-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Baker, Hollie L.
; REGISTRATION NUMBER: 31,321
; REFERENCE/DOCKET NUMBER: 102286.377
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 526-6000
; TELEFAX: (617) 526-5000
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1663 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-793-126-1

Query Match 100.0%; Score 8609; DB 2; Length 1663;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGPTSGPSLLLLLLTHPLA	GNYSITTNILRLSEETWLEADAGDVPVTVH	60
Db	1	MGPTSGPSLLLLLLTHPLA	GNYSITTNILRLSEETWLEADAGDVPVTVH	60
Qy	61	DFPGKLVLSSEKTVLTPATNMGNVTFTIPANREFKSEKGRNKFVTQATFGQVVEKV	120	

Db 61 DFPKGLVLSSEKTLTPATNMGVNTFTIPANREFKSEKGRNKFVTVQATFGVVEKV 120
Qy 121 VLVSLOSYLFIQDXTIYTPGSTVLYRIFTVNHKLLPVGRVWVNIENPESGIPVKQDSL 180
Db 121 VLVSLOSYLFIQDXTIYTPGSTVLYRIFTVNHKLLPVGRVWVNIENPESGIPVKQDSL 180
Qy 181 SSQNLGVLPLSWDIPELVNMGMKIRAYENSPQCVFSTEFVEKEYVLPSEFVIVEPTE 240
Db 181 SSQNLGVLPLSWDIPELVNMGMKIRAYENSPQCVFSTEFVEKEYVLPSEFVIVEPTE 240
Qy 241 KFYIYNEKLEVTITATFLYKKGVEGTAFVIFIGDGEORTSLPESLKRPIEDGSGEV 300
Db 241 KFYIYNEKLEVTITATFLYKKGVEGTAFVIFIGDGEORTSLPESLKRPIEDGSGEV 300
Qy 301 VLGRKVLDDGVQNPRAEDLVGSLYVSATVILHSGSDMVQABRSIGIPVTSPIQHFTKT 360
Db 301 VLGRKVLDDGVQNPRAEDLVGSLYVSATVILHSGSDMVQABRSIGIPVTSPIQHFTKT 360
Qy 361 PKYFKGMPDLVFTNPDGSPAYRVPVAVQGEDTVQSLTQDGVAKLSINTHPSOKPL 420
Db 361 PKYFKGMPDLVFTNPDGSPAYRVPVAVQGEDTVQSLTQDGVAKLSINTHPSOKPL 420
Qy 421 SITVRTKQBSAEQATRTMQALPYSTVGNSSNNYHLVSRTLRPGETLNVLFLRMD 480
Db 421 SITVRTKQBSAEQATRTMQALPYSTVGNSSNNYHLVSRTLRPGETLNVLFLRMD 480
Qy 481 RAHEAKIRYTYLIMNKGRLLAGQVRPQDQVLPVLSITTFIPISFRVAVYTLIGA 540
Db 481 RAHEAKIRYTYLIMNKGRLLAGQVRPQDQVLPVLSITTFIPISFRVAVYTLIGA 540
Qy 541 SQBEVADVWVNDVDSVCSLVKSGSESDRQVPGQMTLKLIEGDHGRVVLVAVDK 600
Db 541 SQBEVADVWVNDVDSVCSLVKSGSESDRQVPGQMTLKLIEGDHGRVVLVAVDK 600
Qy 601 GVFVINKNKLITQSKIWDVVEKADICTPGSGDKYAGVFSAGLTFITSSSQQTQARAEL 660
Db 601 GVFVINKNKLITQSKIWDVVEKADICTPGSGDKYAGVFSAGLTFITSSSQQTQARAEL 660
Qy 661 QCPQARRRSVQLTEKMDKVGKPKELRCCDGMRENPMRPFSCORRTFISLGEAC 720
Db 661 QCPQARRRSVQLTEKMDKVGKPKELRCCDGMRENPMRPFSCORRTFISLGEAC 720
Qy 721 KKVFLDCNVTITELRQHARASHGLARSNLDEIIAENIVSRSEFESLWNLVDELKE 780
Db 721 KKVFLDCNVTITELRQHARASHGLARSNLDEIIAENIVSRSEFESLWNLVDELKE 780
Qy 781 PKQNGISTKLMNIFLKDSTITWELLAVNSDKKICVADPPEVTVMDPFDLRLPYSV 840
Db 781 PKQNGISTKLMNIFLKDSTITWELLAVNSDKKICVADPPEVTVMDPFDLRLPYSV 840
Qy 841 RNEQVEIRAVLYNQELKVRVLLHNPACSLATTKRRHQQTITIPPKSSLSVPYVI 900
Db 841 RNEQVEIRAVLYNQELKVRVLLHNPACSLATTKRRHQQTITIPPKSSLSVPYVI 900
Qy 901 VPLKTGLQVEVKAAYVHFISDVGRKSLVVPVGRVNVKTVAVRTLPDLREGVQKE 960
Db 901 VPLKTGLQVEVKAAYVHFISDVGRKSLVVPVGRVNVKTVAVRTLPDLREGVQKE 960
Qy 961 DIPPADLSQVPTESRILQGTPTVAQMTEDAVDAERLKLIVTPSGCGEQNMIGWTP 1020
Db 961 DIPPADLSQVPTESRILQGTPTVAQMTEDAVDAERLKLIVTPSGCGEQNMIGWTP 1020
Qy 1021 TVIAVHYLDTEQWEXFLEKRGALGLIKGYTQQLAFROPSSAFKVRAPSTWLT 1080
Db 1021 TVIAVHYLDTEQWEXFLEKRGALGLIKGYTQQLAFROPSSAFKVRAPSTWLT 1080
Qy 1081 YVVKVFLAVNLIAIDSVLCGNVKNLILEKQKPGVQEDAPVHOEMIGGLNNNEKD 1140
Db 1081 YVVKVFLAVNLIAIDSVLCGNVKNLILEKQKPGVQEDAPVHOEMIGGLNNNEKD 1140
Qy 1141 MALTAFLVLSIQBAKICEQVNSLPGSITKAGDFLEANNYMLQRSYTVTAIAGYALQMG 1200

Db 1141 MALTAFLVLSIQBAKICEQVNSLPGSITKAGDFLEANNYMLQRSYTVTAIAGYALQMG 1200
Qy 1201 RLKGPLNKLFTTAKDKNRWEDFGKOLYNVEATSYALLALLQKDFDVPVVRWLNQOR 1260
Db 1201 RLKGPLNKLFTTAKDKNRWEDFGKOLYNVEATSYALLALLQKDFDVPVVRWLNQOR 1260
Qy 1261 YYGSGYSTQATFMVFOALAQYOKDAPDHOELNVLVSQLPSSSKITHRIHWSASLLR 1320
Db 1261 YYGSGYSTQATFMVFOALAQYOKDAPDHOELNVLVSQLPSSSKITHRIHWSASLLR 1320
Qy 1321 SEETKENEGFTVTAEGKGQGLSVWVTHYHAKAKDQLTCKNFKDLKVTIKPAPETEKRPDA 1380
Db 1321 SEETKENEGFTVTAEGKGQGLSVWVTHYHAKAKDQLTCKNFKDLKVTIKPAPETEKRPDA 1380
Qy 1381 KNTWILICITRYRGDQATWSILDI SMVTGFPADTDDDLKQLANGVDRYISKYELDKAFSD 1440
Db 1381 KNTWILICITRYRGDQATWSILDI SMVTGFPADTDDDLKQLANGVDRYISKYELDKAFSD 1440
Qy 1441 RNTLIIIVLDKXVSHSEDDCLAFKVHQVFNVELIQPGAVKYVAYYNLEESCTRFYHPEKEDG 1500
Db 1441 RNTLIIIVLDKXVSHSEDDCLAFKVHQVFNVELIQPGAVKYVAYYNLEESCTRFYHPEKEDG 1500
Qy 1501 KNLKCLDELCRCABENCFIQKSDDKVTLEERLDKACBPGVDYVYKTRLVKQVLSNDPDE 1560
Db 1501 KNLKCLDELCRCABENCFIQKSDDKVTLEERLDKACBPGVDYVYKTRLVKQVLSNDPDE 1560
Qy 1561 YIMAETIKSGSDEVQVGGQRTFISPIKREALKLEEKHYLMWGLSSDFWGEKPNLSY 1620
Db 1561 YIMAETIKSGSDEVQVGGQRTFISPIKREALKLEEKHYLMWGLSSDFWGEKPNLSY 1620
Qy 1621 IIGKDTVWEHWPEDBCQDEENKOCODLGAFTESSWVFGCPN 1663
Db 1621 IIGKDTVWEHWPEDBCQDEENKOCODLGAFTESSWVFGCPN 1663

RESULT 2

US-09-132-271-1
Sequence 1, Application US/09132271
Patent No. 6221657
GENERAL INFORMATION:
APPLICANT: Harrison, Richard Alexander
APPLICANT: Faries, Charles Timothy
TITLE OF INVENTION: MODIFIED HUMAN C3 PROTEINS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: HALE AND DORR LLP
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: United States of America
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/132,271
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/793,126
FILING DATE: 07-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Baker, Hollie L.
REGISTRATION NUMBER: 31,321
REFERENCE/DOCKET NUMBER: 102286.377
TELEPHONE: (617) 526-6000
TELEFAX: (617) 526-5000
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1663 amino acids

; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-09-132-271-1

Query Match 100.0%; Score 8609; DB 3; Length 1663;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGPTSGPSLLILLTHPLALGSPWYSIIITENILRLSEBETMYLEADHDAQGVPTVTVH 60
 DB 1 MGPTSGPSLLILLTHPLALGSPWYSIIITENILRLSEBETMYLEADHDAQGVPTVTVH 60
 QY 61 DFPKXVLSSSEKTVLPATNMGNVFTTIPANREPKSEKGRNKFYVQATGTQVVEKV 120
 DB 61 DFPKXVLSSSEKTVLPATNMGNVFTTIPANREPKSEKGRNKFYVQATGTQVVEKV 120
 QY 121 VLVSLSQSGVLFOTDKTITPGSTVLYRIFTVNHKLLPVGRTVVNNIENPEGIPVKQDSL 180
 DB 121 VLVSLSQSGVLFOTDKTITPGSTVLYRIFTVNHKLLPVGRTVVNNIENPEGIPVKQDSL 180
 QY 181 SSQNLGVLPISWDIPELVNNQWKNIRAYYENSPOQVFSTPEPVKBYVLPSEVIVEPTE 240
 DB 181 SSQNLGVLPISWDIPELVNNQWKNIRAYYENSPOQVFSTPEPVKBYVLPSEVIVEPTE 240
 QY 241 KFYIYNEKGLVITITARFLYKKGVEGTAFFVIGIQDGEORISLPSLXRIPIEDSGSEV 300
 DB 241 KFYIYNEKGLVITITARFLYKKGVEGTAFFVIGIQDGEORISLPSLXRIPIEDSGSEV 300
 QY 301 VLRSKVLIDGVONPRAEDLVGKSLYVSATVILHSGSDMVQAESGPIVTSYQHFHTKT 360
 DB 301 VLRSKVLIDGVONPRAEDLVGKSLYVSATVILHSGSDMVQAESGPIVTSYQHFHTKT 360
 QY 361 PKYFKPGMPFDLMVFVTPNDCGSPAYRVPVAVQGEDTVQSLTQGDGVAKLSINTHESQKPL 420
 DB 361 PKYFKPGMPFDLMVFVTPNDCGSPAYRVPVAVQGEDTVQSLTQGDGVAKLSINTHESQKPL 420
 QY 421 SITVTYKQELSEAEQARTWQALPYSTVGNNSNYHLVRLTELPGETLVNHLRLMD 480
 DB 421 SITVTYKQELSEAEQARTWQALPYSTVGNNSNYHLVRLTELPGETLVNHLRLMD 480
 QY 481 RAHEAKIRYTYTILMNKGRLLKAGQVREPQDLVLPISLTITDFIPSPRLVAYVTLIGA 540
 DB 481 RAHEAKIRYTYTILMNKGRLLKAGQVREPQDLVLPISLTITDFIPSPRLVAYVTLIGA 540
 QY 541 SGREVVADSVVWVVKDCVGLVYKSGQSDRQVPGQOQWTLKIEGDHGARVLYAVDX 600
 DB 541 SGREVVADSVVWVVKDCVGLVYKSGQSDRQVPGQOQWTLKIEGDHGARVLYAVDX 600
 QY 601 GFVNLKXNKLTQSKINDVWEKADIGCTPGSGKDYGAVFSDAGLTFTSSGGQQTARABL 660
 DB 601 GFVNLKXNKLTQSKINDVWEKADIGCTPGSGKDYGAVFSDAGLTFTSSGGQQTARABL 660
 QY 661 QCQPAARRRSVOLTEKMDKVGKPKELRKCCEDGWRNPNRPFSCQRTSFISLGEAC 720
 DB 661 QCQPAARRRSVOLTEKMDKVGKPKELRKCCEDGWRNPNRPFSCQRTSFISLGEAC 720
 QY 721 KKVFLDCCNYITELRRQCHARASHLGLARSLNDEIIAENIVSRSEFPESWLNVEDLKE 780
 DB 721 KKVFLDCCNYITELRRQCHARASHLGLARSLNDEIIAENIVSRSEFPESWLNVEDLKE 780
 QY 781 PPKNGISTKLMMNIFLKDSITTWELAVMSDKKIGICVADPFVTVMQDFFIDRLPYSVV 840
 DB 781 PPKNGISTKLMMNIFLKDSITTWELAVMSDKKIGICVADPFVTVMQDFFIDRLPYSVV 840
 QY 841 RNEQVEIRAVLYNFRQNELKVRVELLHNPAFCSLATTKRHOOTITIPKSSLSVPYVI 900
 DB 841 RNEQVEIRAVLYNFRQNELKVRVELLHNPAFCSLATTKRHOOTITIPKSSLSVPYVI 900
 QY 901 VPLKTLGLOVEVKAAYHHFISDGVKRSKLVKVPPIRMNKTAVRVLDPERLGREVOKE 960
 DB 901 VPLKTLGLOVEVKAAYHHFISDGVKRSKLVKVPPIRMNKTAVRVLDPERLGREVOKE 960

QY 961 DIPPADLSQVPTSETRILLQCTPVAQMTEDAVDAERLKHLLVTPSGGCEQNMIGTTP 1020
 DB 961 DIPPADLSQVPTSETRILLQCTPVAQMTEDAVDAERLKHLLVTPSGGCEQNMIGTTP 1020
 QY 1021 TVIAVHYLDETEQWEKFGLEKQGALELIKKGYTQQLAFRQPSAFAAFVKRAPSTWLTA 1080
 DB 1021 TVIAVHYLDETEQWEKFGLEKQGALELIKKGYTQQLAFRQPSAFAAFVKRAPSTWLTA 1080
 QY 1081 YVVKVESLAVNLIAIDSQLCGAVKWLILEKQKPGVFOEDAPVTHOEMIGGLRNNNEKD 1140
 DB 1081 YVVKVESLAVNLIAIDSQLCGAVKWLILEKQKPGVFOEDAPVTHOEMIGGLRNNNEKD 1140
 QY 1141 MALTAFLVLSLOEAKDICEEQVNSLPFSITKAGDFLEANYNMLQSRVTVIAAGVALAQM 1200
 DB 1141 MALTAFLVLSLOEAKDICEEQVNSLPFSITKAGDFLEANYNMLQSRVTVIAAGVALAQM 1200
 QY 1201 RLKGPLLNKFLITAKDKRWEDPGQLXNVEATSVALLIQLKDFDPVVPVRLNEOR 1260
 DB 1201 RLKGPLLNKFLITAKDKRWEDPGQLXNVEATSVALLIQLKDFDPVVPVRLNEOR 1260
 QY 1261 YGGGYGSGTQATFMVFPQALAQYQKQADPDHOBELNDVSLQPSRSKITHRIHWEASLLR 1320
 DB 1261 YGGGYGSGTQATFMVFPQALAQYQKQADPDHOBELNDVSLQPSRSKITHRIHWEASLLR 1320
 QY 1321 SEETKENEFTVTAEKGGQGTLSVVVTMYHAKAKQDLCNKFDLKVTTIKPAPETEKRPQDA 1380
 DB 1321 SEETKENEFTVTAEKGGQGTLSVVVTMYHAKAKQDLCNKFDLKVTTIKPAPETEKRPQDA 1380
 QY 1381 KNTMILEICTRVRGODATMSILDISMNTGAPDTDDLKQLANGVDRIYSKYELDKAFSD 1440
 DB 1381 KNTMILEICTRVRGODATMSILDISMNTGAPDTDDLKQLANGVDRIYSKYELDKAFSD 1440
 QY 1441 RNTLIIYLDKVSHEDDCLAFKVOHYFNVELIQGAVKVAAYYNLESCTRFVHPEKEDG 1500
 DB 1441 RNTLIIYLDKVSHEDDCLAFKVOHYFNVELIQGAVKVAAYYNLESCTRFVHPEKEDG 1500
 QY 1501 KLKLCRDLCELCABENCFIQSDDKVTLERLDKACEPGVDVYVYKRLVKVQUSNDFDE 1560
 DB 1501 KLKLCRDLCELCABENCFIQSDDKVTLERLDKACEPGVDVYVYKRLVKVQUSNDFDE 1560
 QY 1561 YIMAEQTIKSGSDVQVQQTFTFISPKCREALKLEEKHYLMWGLSSDFWGEKPNLSY 1620
 DB 1561 YIMAEQTIKSGSDVQVQQTFTFISPKCREALKLEEKHYLMWGLSSDFWGEKPNLSY 1620
 QY 1621 IIGKDTWVHEPDEECODEENKQKQDLAGFTESWVVFQCPN 1663
 DB 1621 IIGKDTWVHEPDEECODEENKQKQDLAGFTESWVVFQCPN 1663

RESULT 3

US-09-142-334-22
 ; Sequence 22, Application US/09142334
 ; Patent No. 6265485
 ; GENERAL INFORMATION:
 ; APPLICANT: Paries, Timothy C.
 ; APPLICANT: Harrison, Richard A.
 ; TITLE OF INVENTION: Down-Regulation Resistant C3 Convertase
 ; FILE REFERENCE: 4-30443/A/IMU/PCT
 ; CURRENT APPLICATION NUMBER: US/09/142,334
 ; CURRENT FILING DATE: 1999-04-15
 ; EARLIER APPLICATION NUMBER: PCT/CB97/00603
 ; EARLIER FILING DATE: 1997-03-04
 ; NUMBER OF SEQ ID NOS: 35
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 22
 ; LENGTH: 1663
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-142-334-22

Query Match 100.0%; Score 8609; DB 3; Length 1663;
 Best Local Similarity 100.0%; Pred. No. 0;

Matches 1663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	MGPTSGPSSLLLLLLTHPLALGSPMYSIITPNILRESEETWVLEAHDAQDQVPTVTVH	60
DB	1	MGPTSGPSSLLLLLLTHPLALGSPMYSIITPNILRESEETWVLEAHDAQDQVPTVTVH	60
QY	61	DFPCKKVLVSEKTVLTPATNHNMGVTFPIIPANREFEKSEKGNKEVTVQATFGTVVEKV	120
DB	61	DFPCKKVLVSEKTVLTPATNHNMGVTFPIIPANREFEKSEKGNKEVTVQATFGTVVEKV	120
QY	121	VLVSLQGYLFIQDXTIYTPGSTVLVRIFTVNHKLLPVRTVMVNIENPEGIPVKQDSL	180
DB	121	VLVSLQGYLFIQDXTIYTPGSTVLVRIFTVNHKLLPVRTVMVNIENPEGIPVKQDSL	180
QY	181	SSQNQLGVPLSWDIPELVNMGWKIRAYYENSPOQVSTSEFEVKEYVLPSEFIVEPTE	240
DB	181	SSQNQLGVPLSWDIPELVNMGWKIRAYYENSPOQVSTSEFEVKEYVLPSEFIVEPTE	240
QY	241	KFYIYINKEGLEVTITAREFLYKKGVEGTAFVIFGIQDGEQRIISLPESLKRPIEDGSGEV	300
DB	241	KFYIYINKEGLEVTITAREFLYKKGVEGTAFVIFGIQDGEQRIISLPESLKRPIEDGSGEV	300
QY	301	VLRSKVLDDGVQNPRAEDLVGKSLVYSATVILHSGSDMVQAPRSQIPVTSPIYQIHFTKT	360
DB	301	VLRSKVLDDGVQNPRAEDLVGKSLVYSATVILHSGSDMVQAPRSQIPVTSPIYQIHFTKT	360
QY	361	PKYFKPGMPDLVFNTPNPDGSPAYRVPVAVOGEDTVQSLTQDGVAKLSINTHPSOKPL	420
DB	361	PKYFKPGMPDLVFNTPNPDGSPAYRVPVAVOGEDTVQSLTQDGVAKLSINTHPSOKPL	420
QY	421	SITVTRTKQBELSAEQATRTMQALPYSTVGNSSNNVILHSLVRLTELPGETLNVNFFLRMD	480
DB	421	SITVTRTKQBELSAEQATRTMQALPYSTVGNSSNNVILHSLVRLTELPGETLNVNFFLRMD	480
QY	481	RAHEAKIRYTYLIMNKGRLKAGROVREPQGLVPLSITDPIPSFRVAVAYTLIGA	540
DB	481	RAHEAKIRYTYLIMNKGRLKAGROVREPQGLVPLSITDPIPSFRVAVAYTLIGA	540
QY	541	SGQREVVDVWVDKSDCVGSLVWKSQSEDRQVPVQGMFLKIEGHDGARVVLVAVDK	600
DB	541	SGQREVVDVWVDKSDCVGSLVWKSQSEDRQVPVQGMFLKIEGHDGARVVLVAVDK	600
QY	601	GVFLVKNKLTOSKTDVWVEKADIGCTPGSGKDAGVPSDAGLTFTSSSQCOQAORAE	660
DB	601	GVFLVKNKLTOSKTDVWVEKADIGCTPGSGKDAGVPSDAGLTFTSSSQCOQAORAE	660
QY	661	QCQOPAAARRRSVQLTEKRMKDKYKPKELRKCCEGDMRENPMRFSQORRTFRFISLGEAC	720
DB	661	QCQOPAAARRRSVQLTEKRMKDKYKPKELRKCCEGDMRENPMRFSQORRTFRFISLGEAC	720
QY	721	KVFLDCCNVTILRRQHARASHGLHARSNLDEIIARENIVSRSEPPESLWNVDELKE	780
DB	721	KVFLDCCNVTILRRQHARASHGLHARSNLDEIIARENIVSRSEPPESLWNVDELKE	780
QY	781	PPKNGISTKLNPLFKDSITTEILAVSDSKKICVADPEVTVMQDFFIDLRLPSVY	840
DB	781	PPKNGISTKLNPLFKDSITTEILAVSDSKKICVADPEVTVMQDFFIDLRLPSVY	840
QY	841	RNEQVEIRAVLYNVRQNELKRVLELHNPFCSLATTKRRHQQTITIPPKSSLSVPYVI	900
DB	841	RNEQVEIRAVLYNVRQNELKRVLELHNPFCSLATTKRRHQQTITIPPKSSLSVPYVI	900
QY	901	VPLKGTQOEVEVAAVYHHFISQVKSILKVPGEIRMNKTIVAVRTLDPERLREGVQKE	960
DB	901	VPLKGTQOEVEVAAVYHHFISQVKSILKVPGEIRMNKTIVAVRTLDPERLREGVQKE	960
QY	961	DIPPADLSQVPTTESETRILOQTEPVAQMTEDAVDAERLKHILVTPSGCGEQNMIGWTP	1020
DB	961	DIPPADLSQVPTTESETRILOQTEPVAQMTEDAVDAERLKHILVTPSGCGEQNMIGWTP	1020
QY	1021	TVIAVHVLDETEQWEKFGLEKRGALIELIKGYTQOLAFROPSSAPAFVRAESTWLTA	1080
DB	1021	TVIAVHVLDETEQWEKFGLEKRGALIELIKGYTQOLAFROPSSAPAFVRAESTWLTA	1080

QY	1081	YVVKVFLAVNLIAISQVLCGAVKWLILEKQKPDGVFOEDAPVTHQEMIGLRRNNKED	1140
DB	1081	YVVKVFLAVNLIAISQVLCGAVKWLILEKQKPDGVFOEDAPVTHQEMIGLRRNNKED	1140
QY	1141	MALTAFLVLSIQBAKDIQCEQVNSLPGSITTKAGDFFLEANYMNLQRSYTVATAGVALAQMG	1200
DB	1141	MALTAFLVLSIQBAKDIQCEQVNSLPGSITTKAGDFFLEANYMNLQRSYTVATAGVALAQMG	1200
QY	1201	RLKGPLANKELTTAKOKRWEDPGKOLYNVEATSVALLALLOKDFDFPVPVWRLNEQR	1260
DB	1201	RLKGPLANKELTTAKOKRWEDPGKOLYNVEATSVALLALLOKDFDFPVPVWRLNEQR	1260
QY	1261	YVGGYGSTQATFMVFOALAQYQKADPHQELNLDVSLQPSRSSKITTHRIHWESASLLR	1320
DB	1261	YVGGYGSTQATFMVFOALAQYQKADPHQELNLDVSLQPSRSSKITTHRIHWESASLLR	1320
QY	1321	SEETKEGEGTVAEGKGGOTLSVWTHYHAKAKDQLTCKNFKLVTKIYDAPETERPQDA	1380
DB	1321	SEETKEGEGTVAEGKGGOTLSVWTHYHAKAKDQLTCKNFKLVTKIYDAPETERPQDA	1380
QY	1381	KNTMILEICTRYRGDQDATMSILDISMTGTFAPDQDQKQLANGVDRIYSKYELDKAFSD	1440
DB	1381	KNTMILEICTRYRGDQDATMSILDISMTGTFAPDQDQKQLANGVDRIYSKYELDKAFSD	1440
QY	1441	RNTLIIYLDKVSSEDDCLAFKVOYFNVELICPGAVKIVAYNLEESCTFHYHEKEDG	1500
DB	1441	RNTLIIYLDKVSSEDDCLAFKVOYFNVELICPGAVKIVAYNLEESCTFHYHEKEDG	1500
QY	1501	KLKLCDLRCRAEENCFIQKSDDKVTLEERLDKACBPGVDYVYKTRLVQVLSNDFDE	1560
DB	1501	KLKLCDLRCRAEENCFIQKSDDKVTLEERLDKACBPGVDYVYKTRLVQVLSNDFDE	1560
QY	1561	YIMAIBOTIKSGDEVOVGQORTFISPIKCRALKEKKHYLMWGLSSDFWGEKPNLSY	1620
DB	1561	YIMAIBOTIKSGDEVOVGQORTFISPIKCRALKEKKHYLMWGLSSDFWGEKPNLSY	1620
QY	1621	IIGKDTVWEHPBEDECCQDEENKQCODLGAFTESWVFGCPN	1663
DB	1621	IIGKDTVWEHPBEDECCQDEENKQCODLGAFTESWVFGCPN	1663

RESULT 4
US-08-447-411-2
Sequence 2, Application US/08447411
Patent No. 5773243
GENERAL INFORMATION:
APPLICANT: FRITZINGER, DAVID C.
APPLICANT: BREDEHORST, REINHARD
APPLICANT: VOGEL, CARL-WILHELM
TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,411
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/043,747
FILING DATE: 07-APR-1993
ATTORNEY/AGENT INFORMATION:

NAME: Oblon, No. 5773243man F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 1126-101-0
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 413-3000
 TELEFAX: (703) 413-2220
 TELEX: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1651 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-447-411-2

Query Match 51.0%; Score 4389.5; DB 1; Length 1651;
 Best Local Similarity 51.3%; Pred. No. 0;
 Matches 853; Conservative 309; Mismatches 471; Indels 31; Gaps 18;

9 LLLLLLTHLPALGSPMYSIIITNIRLSEETMVLAEHDAOGDVPVTVTHDFPGKLV 68
 8 LVAALLIGFGSGHGYLTITPAVURTTEQILVEAGDSTPKSLDIFVHDFPRKQT 67
 69 LSSEKTVLTATNMGNVTTIPANKFEKSEKGRNKFTVQATFTQV-VEKVLVLSQS 127
 68 LFQSRVDMNAGSNEFTPTIKVPA-KELNKSQNCQVYVVKVT-GPOVALEKVVLLSYQS 125
 128 GYLPIQDKTIYTPGSTVLRYIFTVNHKLLPVGRITVMVNIENPEGIPVKQDSLSQNLG 187
 126 GFVPIQDKGIYFGSPVRVRSVDNHRMDXIVIVEFCQPEGLVVSKEFVPSGSI- 184
 188 VLPLSWDIPELVNGQWKIBAYVENSQQVFSTEFVEKYYVLPSEVIVPEPEKYYIYN 247
 185 ---RPNLPLSVSGFTKAVAKYEHSPESYATAYDFVREYVLPSEVRLQPSDKFLYIDG 241
 248 EKGLEVTITARFLYGGKVEGTAFVIGIODGEORISLPSLAKIPIEDSGSGWLSRKVL 307
 242 NKPHVSITARYLYGKVEGVAVGVGKLDKAKSIPDSLRIFIIDGGEATLKRTL 301
 308 LDGVNPRABDLVKSLYVATVILHSGSDMQAERSGIPVTSYQIHFTKPYKFG 367
 302 RSREPD--LNQLVGHLYVSVTVITBESGDMVVTBEGGHIYVTSYQIYFTKPYKFG 359
 368 MPDLMVFTNPDGSPAYRPV--AVQGEDTVQSLTQGDGVAKLSINTHPSQKPLSTV 424
 360 MPYELTVYVNPDGSPAAHVVPVSEAHSEGT---TLDGTAKLINTPLNIQSLPITV 415
 425 RTKQELSEAEQATRTMQLPVPYTVGNSNNYHLVRLTELPGETLVNVLNLRMDRAHE 484
 416 RTNHGDLPRERQAIKSMATAYQTQGSSENYLHVAITSTETKPGDNLPLVFNVRGNANSL 475
 485 AKIRYVYTLWKNKGRLLKAGQVREPGQDLVLPISITTDFTPSFELVAYTYLICASGR 544
 476 NQIKFYIYLINLKGKIFKVGQRPDRDQNLVTLNHLITPDLPSPFRVAYYQV---GN 531
 545 EVVADSVVDVKSCVGLVVKSGS-EDRPVFGQOMTLKIEGHBGARVVLVADKGVF 603
 532 EIVADSVVDVKDTCMGTLLVVGASSRDRRIQKFGAAMKIKLEGDPGARVGLVAVDKAVY 591
 604 VLNKNKLTOSKIDWVEKADIGCTPGSGKDYGAVFSDAGLFTSSSGQQTAAQRAELQCP 663
 592 VLNDKYKISQAKIWDITIEKSDFGCTAGSGQNLGVEDAGLALTSTNLNLTQKRAAKCP 651
 664 QPAARRR-SVLTEKMDKVGKYPKE-LRKCCEDGMENPNRPFSCQRTRPISIGEACK 721
 652 QPANRRRSSVLLDSSKASAAQFQDQGLEKCCEDGMENPNMGYTCEKRAKYIQBGDAK 711
 722 KVFLLDCNNTYELRQARASHLGLARSLNLDIIAEENIVSRSPFESLWNNVEDLK-E 780
 712 AAFLECCHYIKGIEDENCRSELFLARSDFEDELFGDDNIISRSDFPESLWLTETLGE 771
 781 PPKNGISTKLMNPLKDSITTWILAVMSDKKIGCVADPFVTVMQDFFIDRLRPYSV 840

772 PNNQGISSKTVPFYLRDSITTWELLAVGLSPTKGVCAEPYEITVMKDDFFIDRLRPYSV 831
 841 RNEQVEIRAVLYNRQONELKVRVELLENPAFCSLATTKRRHQOTITIPPKSSLSVPYVI 900
 832 KNEQVEIRAILYNY-ADEDIYVRVELIYNPAFCASASTEGORYRQFPKALSSRAVPFVI 890
 901 VPLKTLQEVVEKAAVYHHFISDGVKSLKVVPEGIRNMKTVAVRTLDPELRGEGVQKE 960
 891 VPLEQGLHDEVEIASVRGELASDGVKRLKVVPEGERKN-IVTIELDPSVKGVGQTQEL 949
 961 DIPPADLSQVDPDTESETRILLQGPVPAQMTEDAVDAERLKLHIVTSPGCGEQNMIGWTP 1020
 950 TVIANKLDDKVPDTEVETRIISVLGDPVAQIENSIDGSKLNLHIIITSPGCGEQNMIGWTP 1009
 1021 TVIAVHYLDETEQWEKFGLEKQKGALELIKKGYTOQLAPROPSSAPAFVVKRASTWLT 1080
 1010 SVIATYILDATQWENLGVDRTEAKQIMTGVAQOMVYKKAHSAFTNRASSSLWLT 1069
 1081 YVVKVFLAVNLIIA-IDSOVLGAVKWLILEKQKPDGVQFQEDAPVHQMIGGLRNNEK 1139
 1070 YVVKVLAMASNVMKDISHEIICGVKWLILNRQPDGVFKENAPVHGEMLGGTK-GAEP 1128
 1140 DMALTFVLISLOEAKDICEBOVNSLPGSITKAGDFLEANTMNLQSYTVVAIAGYALAQ 1199
 1129 EASLTAFIVTALLERSVCKEQINTLDSINKATDYLKKYEKLQRPYTTALTAYALAA 1188
 1200 GRLLKGPLLNKFLTTAKDKNRWEDPGKQLYNVEATSYALLALLQLKDPDFVPPVVRWLN 1259
 1189 DLANDDEV--LMAASTGRNNEEYNARTHTNIGTSYALLALLKMKKPAEVPVVRWLDQ 1246
 1260 RYGGGYGTOATFVFOALAQYKXADPHQELNLDVSLQLPSSRSKITHRIHWESALL 1319
 1247 KYGGTYGTOATFVFOALAEYEIQMPTHQDLNLDISIKLPEREVPERYSINDRNAVQA 1306
 1320 RSEETKENGSGFTVTAEGKGGTSLSVVTVMYHAKAK-DQLTCNKFDLKVYTIKPAETKRPQ 1378
 1307 RTVETKLNEDFTVSASGDKATMTILTVYNAQLREDANVCNKFHLDVSVENVELNLQAK 1366
 1379 DAKNWMILBEICTRYRGDQDATMSILDISMTGFPAPDPTDQLKLANGVDVRYISKYELDKAF 1438
 1367 GGKAAALRLKICTRYLGEVDSTMTIISMLTGFPDAEDLKRLENGVDVRYISKYELDNNM 1426
 1439 SDANTLIIIXLDKSHSEDDCLAFKQVQFNVNVELOPGAVKYVAYVYNNLEESCTRFVHPEKE 1498
 1427 AQKGTWVYLDKSHSEDECLHFKHKEFVGFIQPGSVKYVSYNNLEDEQCTKEVHPDKE 1486
 1499 DGKLNKLRDELRCBAENCFIQKSDSKVTLEERLDKACEFGVDYVYKTRLVKQLSNDP 1558
 1487 TGLNKKICHGNTCRCAEETCSLLNQKKIDILQIRIQKACQNVYVYKTKLRLREEKDG 1546
 1559 DEVIMAETQITKSGSDEVQVGQORTFISPIKCREALKLEKKHYLMGLSDFWGEKPNL 1618
 1547 DIVFMDVLEVIKGGTDRNAQAQARQYVSRKQCEALNKLNDVYLINGLSSDLWPMKDDI 1606
 1619 SYIIGKDTVWEHWPBEDECCDEENKQCCQDILGAFTESMWVFGCP 1662
 1607 SYLITKNTWIERPNWDECCQDEEFQNLCDLDFQAUSNTLTITFGCP 1650

RESULT 5

US-08-662-227-2
 ; Sequence 2, Application US/08662227
 ; Patent No. 5922320
 ; GENERAL INFORMATION:
 ; APPLICANT: VOGEL, CARL-WILHELM
 ; APPLICANT: BREDEHORST, REINHORST
 ; APPLICANT: KOCK, MICHAEL
 ; APPLICANT: FRITZINGER, DAVID
 ; TITLE OF INVENTION: RECOMBINANT PROCVF
 ; NUMBER OF SEQUENCES: 39
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
 ; ADDRESSEE: P.C.

STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
 CITY: ARLINGTON
 STATE: VA
 COUNTRY: USA
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 FILING DATE: 14-JUN-1996
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: OBLON, NORMAN F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 1126-0107-0X
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-413-3000
 TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1642 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-662-227-2

Query Match 49.1%; Score 4224; DB 2; Length 1642;
 Best Local Similarity 49.6%; Pred. No. 0;
 Matches 824; Conservative 324; Mismatches 478; Indels 36; Gaps 20;

Qy	9	LLLLLLTHPLALGSPMYIIITPNILRLSEETVLEAHDAQGVPTVTVHDPGKXIV	68
Db	8	LVAALLIGPGSHGALYTLITPAVLRTDTEBQIILVEAHGSDTPKLDIFVHDPFKQKT	67
Qy	69	LSSEKTVLPATNMGNVFTTIPANREFKSEKGRNKFVTVQATFGTV-VEKVLVLSQS	127
Db	68	LFQTRVDMNPAGGMLVPTIEIPA-KEVSTDQRQNVVVQVTP-GPQVLEKVVLLSYQS	125
Qy	128	GYLFIQDTKTIITPGTSLVLYRFTVNHKLLPVGRTWMIENEPGIPVKQDSLSSQNLG	187
Db	126	SFLFIQDTKGIYTPGSPVLYRFSMDHNTSKNKTIVIVEFQTPGILVSSNSVD-----	179
Qy	188	VLPLSW--DPELVNMGOWKIRAYYENSPQVSTEFPEVKEBYVLPSPFVIEPTKPYFI	245
Db	180	-LNFPPYLPDLVSLGTWRIYAKYHSPEN-YTAYFDVRKYVLPSPFVRLQPSKFFYI	237
Qy	246	YNEKGLVITAFPLCYKKEVGTAFVIGIQDGEORISLPSLXRIPIEDSGSERVLSRK	305
Db	238	DGNEFHSVSTAKLYGEEVEGVAFVLFQKIDDAKSIPLSTRIPIDGDKATLKD	297
Qy	306	VLLDGVQNPRAEDLVGKSLYSATVILHSGSDMVQASRGIPVTSYQIHFTKPYFK	365
Db	298	TFRS--RFPMNLVUGHTLYASVTVMTSGSDMVVTEQSGIHIIVASPYQIHFTKPYFK	355
Qy	366	PGMPEDLVFVTPNPGSPAYRVPVAVOGEDTVQSITQCGGVAKLSINTHPSQPLSITVR	425
Db	356	PGMPYELTVFVTPNPGSPAAHVPVSEAFHSV-GTTLSDGTAKLILNPLNAQSLPTIVR	414
Qy	426	TKSQELSEAFQARTWQALPYSTVGNNSNNYLSLVLRTLPBGTLLNVLFLLRMDRAHEA	485
Db	415	THRGDLPRERQATKSMATAYQTQGGSGNYLHVAITSTEIKFGDNLPNFNVKNANSLK	474
Qy	486	KIRYTYLLMNGKLLKAGQVREPCQDLVLPISITTDFTPSRELVAHYTLIGASQRE	545
Db	475	QIKFYTYLLMNGKIKFVGRQPRDQGNLVTVNLNHLITPDLPSFRFVAYVQV-----GNNE	530
Qy	546	VVADSVVMDVKDS CVGSLVVKSGQSEDRQVPFGQWTLKIEGDHGARVVLVAVDKGVFL	605
Db	531	IVADSVVMDVKOTCMGTLVVK---GDNLIQMGGAAMKIKLEGDRGARVGLVAVDKAVYVL	587

Qy	606	NKQKLTQSKIMDVVKADIGCTPGSGKDYAGVFS DAGLTFTSSSQOQTAQBAELQCPQ	665
Db	588	NDKIKSQAKIMDTIEKSDFGCTAGSGQNNLGVFEDAGLALTSTNLTQKSAKACQCPQ	647
Qy	666	AARRRR-SVOLTEKRMQKVGKY-PKELRKCCEGDMRENPMRFSCQRRTRFISLGSACKKV	723
Db	648	ANRRRRSVLLDSNASKAAAFQDQLRKCCEDWHENPMGYTCBKRAKVIQEGDACKAA	707
Qy	724	FLDCNVITELRQHAPASHLGLARSNLDEDIABENIVSRSEFESMLWNVEDI-KEPP	782
Db	708	FLECCRYIKGVRDENQRESEFLARDNEDGFIADSDIISRSDFKSWMLTRDTEPN	767
Qy	783	KNGISTKLMMIFLKDSITTWELLAVMSDKKIGICVADPEFVTVMQDFFIDELPYSVVRN	842
Db	768	SGIISKTMSFYLRDSITTWVLAWSFTPTKIGICVAEPYEVIRVMKVFFIDLOMPYSVVRN	827
Qy	843	FOVEIRAVLYNRYONQELKYRVELLHNPAFCSLATTKRHHQOTIIPPKSSLSVYVTVP	902
Db	828	EQVEIRAILHNY-VNEDIYVRVELLNPAFCSASTKGQRYRQFPKALSSRAVFEVIVP	886
Qy	903	LKTGLQEVVEKAAVYHHFISDGVKSLKVYPEGIRMNKTAVVRLTDPERLGBEGVCKEDI	962
Db	887	LEQGLHDEVEIKASVOEALWSDGVKXKLKVVEGVQ-KSIVTVIKLDPRAKGVGGTQLEVI	945
Qy	963	PPADLSQVDPDTSETRIILQGTPEAQMTEDAVDAERLAKHLIVTPSGCGEQNMIGMTFTV	1022
Db	946	KARKLDRVPDTELETKIILQGPVPAQITENSIDGSKLNHLIITPSGCGEQNMIMAAV	1005
Qy	1023	IAVHYLDETEQWKEFGLEKQGALEIHKGYTQQLAFROPSSAFAPVYKRASTWLTAYV	1082
Db	1006	IATVYLTOTTEQWETLGNRRTEAVNQIVTGYAQWYKKAADSHSYAFTNRASSWLTAYV	1065
Qy	1083	VKVESLAVNLIA-IDSOVLGAVKWLILEKOPDGVQEDAPVHEQMIIGGLRNNEKDM	1141
Db	1066	VKVFAAKVAVGSHIEICGVVWILLNRQDPDGAFAKENAVPLSGTMOGGIQ-CAEEV	1124
Qy	1142	ALTAFTVISTQAKDICEEOVNSLPSITKAGDFLEANTYNNIORSYTVVAIYALAAQWR	1201
Db	1125	YLTAFILVALLESKTCINDVYNSLDSIKKATWYLLKYEKLQRPYTTALTAYALAAQD	1184
Qy	1202	LKGPLANKFLITAKDKNRWEDPGKLYNVVETSYALLALLQLKDFEPVPPVVRWLEORY	1261
Db	1185	LNDRVLMASSTGRD--HNEEYNAHTNTEGTSYALLALLKMKKDFDQGTPIVWLTQNP	1242
Qy	1262	YGGYGSTQATFWFQALAQYKDAFDHBLNLDVSLQPSRSSKITTHRIHWEASLRS	1321
Db	1243	YGETYGTQATWAFQALAEYEIQMPTKDLNLDITIELPDREVPTRVRYNTENALLART	1302
Qy	1322	EETXENEGFTVTAEGKGGTSLVVVTVMYHAKAKDOL-TCNKFDLKVITIKPAPETEKRPDA	1380
Db	1303	VELKLNODITVTASDGKATWLTITFYNALQKANKVCKFHLNVSVE---NIHLNAMA	1359
Qy	1381	KNTWLICITRYRGDQATWISLIDISMTGFAEDTDLLQLANGVDVRYTSKTELKAFSD	1440
Db	1360	KGALMLKICRYLGEVDSTWTIIDISMLTGLPDAEDTLRLSGVDVRYTSRYEVDNNMAQ	1419
Qy	1441	RNTLIIILDKVSHEDDCLAFKHQVFNVELIQGAVKVYVAYVYNNLEESCTRYHPEKEDG	1500
Db	1420	KVAVIIVLNKVSHEDESLHFILKHFEVGFQPGSVKYSYNNLDEKCTKYHDPKGTG	1479
Qy	1501	KLNLCRDEICRABENCFIQKSDSKVLEERLUDKACEPGVDYVYKTRLVQLQNSDFE	1560
Db	1480	LLNKICIGNVCRACAGETCSLSNLQERIDVLPQIEKACETNVDYVYKTKLLRLEEQDGNDI	1539
Qy	1561	YMAIEOTIKSGSDEVQVQOORTFIPICREALKLEKKHYLMWGLSDFWGEKPNLSY	1620
Db	1540	YVMDVLSVIRKQGTDENPRAKTHOYISQKCOEALNKVNDVYILWGSRLDLPKDKISY	1599
Qy	1621	IIGKDTVVEHWPDEDECODEENOKQODLGAFTESVWVFGCP	1662
Db	1600	ITKNTWIERPHEDECEBEFQKLCDDFAQFSYTUTEFGCP	1641

RESULT 6

US-09-017-947-2
 ; Sequence 2, Application US/09017947
 ; Patent No. 6303754
 ; GENERAL INFORMATION:
 ; APPLICANT: VOGEL, CARL-WILHELM
 ; APPLICANT: BREDEHORST, REINHORST
 ; APPLICANT: KOCK, MICHAEL
 ; APPLICANT: FRITZINGER, DAVID
 ; TITLE OF INVENTION: RECOMBINANT PROCVF
 ; NUMBER OF SEQUENCES: 39
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
 ; ADDRESSEE: P.C.
 ; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
 ; CITY: ARLINGTON
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22202
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/017,947
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/662,227
 ; FILING DATE: 14-JUN-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: OBLON, NORMAN F.
 ; REGISTRATION NUMBER: 24,619
 ; REFERENCE/DOCKET NUMBER: 1126-0107-0X
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 703-413-3000
 ; TELEFAX: 703-413-2220
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1642 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-017-947-2

Query Match 49.1%; Score 4224; DB 3; Length 1642;

Best Local Similarity 45.6%; Pred. No. 0;
 Matches 824; Conservative 324; Mismatches 478; Indels 36; Gaps 20;

QY	9	L L L L L L T H P L A L G S P M Y S I T P N I L R L S E E T W L E A H D A Q D V P V T V T V H D F P G K L V 68
DB	8	L V A A L L I G P G S S H G A L Y T L I T P A V L R T D T E E Q I L V E A H G D S T P K Q L D I F V H D F P R K Q K T 67
QY	59	L S S E K V L T R A T N H M G N V T T I P A N R E F K S E K G E N F V T V O A T F G Q V - V E K V L V S L O S 127
DB	68	L F Q R V D N P A G G M L V T T I E I P A - K E V S T D S R Q N Q V V V Q V T - G F C V L E K V G L S L S Y G S 125
QY	128	G Y L F I Q D K T I Y T P G S T V L Y R I F T V N H K L L F V G R T Y M N I E N P E G I P V K Q D S L S S Q N Q L G 187
DB	126	S F L F I Q D K G I Y T P G S P V L Y R V F G M D N T S K M N K T I V E P Q T P E G I L V S S N S V D - - - - - 179
QY	188	V L P L S W - - D I P E L Y N M G W K I R A Y E N S P Q O V S T E F E V K E Y L P S F E V I V E P T K F Y I 245
DB	180	- L N F W P N L P D L V S L G T W R I V A K Y E S P E N - Y T A F D V R K Y V L P S F V R L Q P S E K F F Y I 237
QY	246	Y N E R G L E V T I T A R F L Y G K Y V E G T A F V I F G I Q D G E Q R I S L P E S L K R I P I D G S G E V V L S R K 305
DB	238	D G N E N F H V S I T A R F L Y G E E V G A F V L F G V K I D D A K S I P D S L T R I P I D G D G K A T L K E D 297
QY	306	V L L O G V Q N P R A E D L V G K S L Y S A T V I L H S G D M V Q A E R S G I P V T S P Y Q I H F T K T P K Y E K 365

DB	298	T F R S - - R F P N L N E L V G H T L Y A S V T V N T E S G S M V V T E Q S G I H I V A S P Y Q I H F T K T P K Y E K 355
QY	366	P G M P F D L M V E V T N P D G S P A Y R V P V A V Q G D H T V Q S L T Q S D G V A K L S I N T H P S O K P L S I T V R 425
DB	356	P G M P E L T V V T N P D G S P A A R V P V S E A F H S M - G I T L S D G T A K L I L N I P L N A Q S L P I T V R 414
QY	426	T K Q E L S E A Q A T R T M Q A L P Y S T V G S N N Y L H S L V R T E L R P G E T I N V N F I L R M D R A H E A 485
DB	415	T N H G D L P R E R Q A T K S M T A I A Y Q F G S G N Y L H V A T I S T E I K R G D N L P V N F N V K G N A S L K 474
QY	486	K I R Y V T L I M K G R L K A G R Q V R P Q D L V L P L S I T T D F I P S F R L Y A Y Y T I L I G A S G O R E 545
DB	475	Q I K Y F T V I L N K I F K V G R P R D Q N I V T M N L H T I P D L I P S F R V A Y Y Q V - - - G N N E 530
QY	546	V V A D S V W V D Y K D S C V G S L V V K S G S E D R O P V P Q Q M T L K I E G D H C A R V V L V A V D K G V L 605
DB	531	I V A D S V W V D V K D T C M G T L V W K - - G N L I Q M P G A M K I K L E G D P C A R V G L V A V D K A V V L 587
QY	606	N K N K L T Q S K I W D V E K A D I G C T P G S G K D Y A G V F S D A G L T F T S S S G Q O T A R A E L Q C P Q P 665
DB	588	N D K Y I S Q A K I M D T I E K S D F G C T A G S G Q N N L G V P E D A G L A L T T S N L N T K Q R S A A K C P Q P 647
QY	666	A A R R R - S V O L T E R M D K V G K Y - P K E L R K C C D G H R E N P R E S C O R R T R F I S L G A C K V 723
DB	648	A N R R R S V L L L D S N A S K A A E F Q D Q L R K C C D V H E N P M G V T C E K R A K Y I Q S G D A C K A A 707
QY	724	F L D C C N Y I T E L R Q H A R A S H L G L A R S N L D E D I I A E N I V S R S E F P E S F W L M N V E D L - K E P P 782
DB	708	F L E C C R Y I K G V R D E N Q R E S E L F L A R D N E D G F I A D S D I I S R S D F P K S M L W L T K D L T E E P N 767
QY	783	K N G I S T K M N I F L K D S I T T W E L L A V S M S D K G I C V A D P E V T M O D F I D L R L P Y S V R N 842
DB	768	S Q G I S S K T M S F Y L R D S I T T W V L A V S F T P T K G I C V A E P Y E I R V M K V F I D I O M P L S V V K N 827
QY	843	E Q V E I R A V L Y N Y R Q N Q E L K V R V E L L H N P A F C S L A T T K R H Q O T I T I P P K S E L S V P Y I V P 902
DB	828	E Q V E I R A I L H N Y - V N E D I V R V E L L Y N P A F C S A S T K G Q S Y R Q Q F I K A L S S R A V P E V I V P 886
QY	903	L K T G L O E V E V K A A V Y H F T S D G V R K S L V K V P E G I R M N K T V A V R T L D P E R L G R E G V K E D I 962
DB	887	L E O G L H D V E I K A S V Q E A L W S D G V R K L K V V P G V Q - K S I V T I V L D P R A K V G G G Q L E V I 945
QY	963	P P A D L S D Q V P D T E S E T R I L L Q G T P V A Q M T E D A V A E R L A K H L I V T P S G C G E Q N M I G M T P V 1022
DB	946	K A R K L D R V P D T E I E T K I I I Q G D P V A Q I E N S I D G S K L N H L I T P S G C G E Q N M I R M A A P V 1005
QY	1023	I A V H Y L D E T O E K F G L E K S Q G A L E L I K G Y T Q O L A F R O P S S A F A F V K R A P S T W L T A Y V 1082
DB	1006	I A T Y I D T T E Q M E T L G I N R T E A V N Q I V T G Y A Q Q V Y K A D H S Y A A F T N R A S S W L T A Y V 1065
QY	1083	V K V F S L A V N L I A - I D S Q V L C G A V K M L I L E K Q K P D G V F O E D A P V I H Q E M I G L R N N N E K D M 1141
DB	1066	V K V F A N A A K M V A G I S H E I I C G V R W L I L N R Q Q P D G A F K E N A P V L S G T M Q G G I Q - G A E E V 1124
QY	1142	A L T A F V L I S Q E A K D I C E E Q V N S L P G S I T K A G D F E A N Y M N L O R S Y T V A I A C Y A L A Q M G R 1201
DB	1125	Y L T A F I L V A L L E S K T C N D Y V N S L O S S I K A T N Y L K K Y E K L Q R P Y T T A L T A Y A L A A D Q 1184
QY	1202	L K G E L N K F L T T A K D K R W E D F G K Q L Y N V E A T S Y A L L A L L Q L K D F D F V P V V R W L N E Q R Y 1261
DB	1185	L N D R V L M A A S T G R D - - H W E E Y N A H T N I E G T S Y A L L A L L K M K T D Q T G P I V R W L T D Q N F 1242
QY	1262	Y G G Y G S T O A T F M V F O A L A Q Y K D A P D H O E I N L D S L O L P S R S S K I T H I H W E S A S L A R S 1321
DB	1243	Y G T Y G O T Q A T V N A F O A L A E Y E I Q M T H K D L N D J I T I E L D F R E V I R I N Y E N A L L A R T 1302
QY	1322	E E T K E N E G F T V T A B G X G Q G T L S V V T M Y H A K A Q D L - T C N K F D L K V T I K P A P E T E R K P Q D A 1380
DB	1303	V E T K L N Q D I T V T A S G D G K A T M I L T F Y N A Q L E K A N C N K P H L N V S V E - - - N I H L N A M G A 1359
QY	1381	K N T W I L E I C T R Y E G D O D A T M S T I L D S M T G E A P D T D L K Q L A N G V D R V I S K Y E L D K A P S D 1440

Db 1360 KGALMLKICTRYLGEVDSWTIIDISMLTGFIPDAEDLTRLKSGVDRVLSRYEYDNNMAQ 1419
 Qy 1441 RNTLIIVLDKVSUSDCLAFKHQVFNVELIQGAVKVAAYNLEESCTRFYHPEXEDG 1500
 Db 1420 KVAVIIYNLVKVSSEDECLFKILKHFVGFQPGSVKVSYSYNLDEKCTKFYHPDKGTG 1479
 Qy 1501 KLNKLGRDELRCRAEBCNFIQKSDDKVTLLESLDRACBPGVDYVYKTLVKVQLSNDPDE 1560
 Db 1480 LNNKICIGNVCRCAGETCSLSNHOERIDVPLQIEKACETNVDYVYKTLLEESDQNDI 1539
 Qy 1561 YWIAEOTIKSGDEYVQOQRTFISPIKCRALKLEBKHYLMGLSSDPWGBKPNLSY 1620
 Db 1540 YVMDVLEVIKQGTDENPRAKTHQYISQRKCOBALNLYNDYLIWGSRLPLTKRISY 1599
 Qy 1621 IIGKDTWHEWPEDECOBENOKOQDLCAGATESMVFVGC 1662
 Db 1600 IITKNTWIERWPEDECOBEEFKQLCDDFAQSYTLTBEFGCP 1641

RESULT 7

US-09-925-442-2

; Sequence 2, Application US/09925442

; Patent No. 6607897

; GENERAL INFORMATION:

; APPLICANT: VOGEL, CARL-WILHELM

; BREDEHORST, REINHORST

; KOCK, MICHAEL

; FRITZINGER, DAVID

; TITLE OF INVENTION: RECOMBINANT PROCVF

; NUMBER OF SEQUENCES: 39

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,

; P.C.

; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY

; CITY: ARLINGTON

; STATE: VA

; COUNTRY: USA

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/925,442

; FILING DATE: 10-Aug-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/017,947

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: OBLON, NORMAN F.

; REGISTRATION NUMBER: 24,618

; REFERENCE/DOCKET NUMBER: 1126-0107-0X

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-413-3000

; TELEFAX: 703-413-2220

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1642 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-925-442-2

Query Match

Best Local Similarity 49.1%; Score 4224; DB 4; Length 1642;

Matches 824; Conservative 324; Mismatches 478; Indels 36; Gaps 20;

Qy 9 LLLLLTHPLALGSMYSIITPNILRESEETVLEADAQGDGVFVTVTYHDFPGKKLV 68

Db 8 LVAALLIGPGSSHGALYTLITPAVLRTDTEEOILYEAHGDSTPKQLDIFVHDFPRKQKT 67
 Qy 69 LSSEKTVLTPATNMGVNTFTIIPANREFKSEKGNKFPVVOATFGTQV-VEKVLVLSLOS 127
 Db 68 LFOQVRDMNPAGGMLVFTTTEIPA-KEVSTDSQNOYVVQVVT-GEQVLEKVKVLLSVO 125
 Qy 128 GYLFTQDCKTYITPGSTVLYRIFTVNHKLLPVQRTVWANIENPEGIPVQDLSLSONQLG 187
 Db 126 SFLFIQTDKGIYITGSPVLYRVFSMDHNTSKMNTVIIVEFQTEGILVSSNVD----- 179
 Qy 188 VLPILGW--DIPELVNMQWKIRAYENSPOOVSTSEFVKEVYLPSEFVIVTEFTEFYI 245
 Db 180 -LNPFWPNLPLDLSLGTWRIVAKYHSPEN-YTAYFDVRKYVLPSEFVRLQFSBKFFYI 237
 Qy 246 YNEKLEVTITARELYGKVEGTAFVIFGIQDGEQISLPESIKRPIPIEDGSGEVVLGRK 305
 Db 238 DGENFHVSIITARYLYGEEVGVAFVIFGVKIDDAKSIIPDSITRPIIDGDKATLKED 297
 Qy 306 VLLDGVQNPRAEDLVKSLVYSATVILHSGSDMVQVQERSGIPVITSFVQIHTKPKYK 365
 Db 238 TFRS--RFPNLNELVGHLYASVTVMTESGSDMVVTEQSGIHVAFYQIHFKTKPKYK 355
 Qy 366 PGMPFDLMVFVTNPDGSPAYRVPVAVQEDTVQSLTQGGQVAKLSINTHPSOKPLSITVR 425
 Db 356 PCMPVELTVYVTNPDGSPAAHVPVVSFAFHM-GTLLSDGTAKLILNIPUNAQSPLITVR 414
 Qy 426 TKQELSEAOATHTMOALPYSTVGNNSYVILHLSVRLTELPGETLVNVLRLMDRAHEA 485
 Db 415 TNHGDLPREOATKSMTAIAYQTQGGSGNYLHVATITBIKPGDNLVFNFNKGNANSUK 474
 Qy 486 KIRYVYVILNMKGLKAGQVREPQDLVVLPLSITTDPIPSFRLLVAYITLIGASGORE 545
 Db 475 QIKYFTVILNKGKIFKVGQPRDQGNLVNMLHITPDLIPSEFVAYYQV---GNNE 530
 Qy 546 VVADSVWVDVKDSCVQSLVNVKSGSEDRQVPVQOQWTLKIEGDHGRVVLVAVDKGVFL 505
 Db 531 IVADSVWVDVKDTCMGTLVVK--GDNLIQMPGAANKIKLEGDPGARVGLVAVDKAVYL 587
 Qy 606 NKQKLTQSKIWDVVEKADIGCTPGSGKVAGVFSAGLTFTSSSQOQTAQRAELQCQOP 665
 Db 598 NDKYKISQAKINDTIEKSDFGCTAGSGONNLGVFEDAGLALTSTNLNFKQSAAKCQOP 647
 Qy 666 AARRRR-SVOLTEKRMKVKY-PKELRKCCEGMBENMRPSCQRTFISLGEACKV 723
 Db 648 ANRRRRSVLLDSDNASKAAEFQODLRKCCEDMBENPMGYTCERAKYIIEGDACKAA 707
 Qy 724 FLDCCNVITELRRQHARASHLGLARSNLDBDIIAENIVSRSEFFESLWNVEDL-KEPP 782
 Db 708 FLECCRYIKGVDRDENQRESELFLARDNEDGFIADSDIISRSDFPKSLWLTGKLTEDN 767
 Qy 793 KNGISTKLMTFLKDSITTHEILAVSMDKGI CVADPPEVTVMODFFTLRLPYSVVN 842
 Db 768 SGISSTKMSPLYRDSITVWVLAFTPTKGCVAEPYRVRWKVFFIDLPQPSVVKXN 827
 Qy 843 EQVEIRAVLYNRQNELKVRVELLHNPAPCSLATTKRSHQQTITIPPKSSLSVPYVVP 902
 Db 828 EQVEIRAILHNY-VNEDIYVRVELLYNPAFCASASTKQRYRQOFPIKALSSRAFPVVP 886
 Qy 903 LKTGLQEVKAAVYHHFISDGVKSLKVYPEGIRMNKTVAVPTLDPRLREGVQKEDI 962
 Db 887 LEQGLHDEVEIKASQOEALNSDGVKKLKVVEGVQ-KSIVTVKLDPRAKGVGGTQLEVI 945
 Qy 963 PPADLSQVPTSETETRIILQGTTPVAQMTEDADAERLKEHIVTPSGQEQNMIMTPTV 1022
 Db 946 KARKLDDRPDTEIETKIIQQDPVAQIIENSIDGSKLNLHIIITPSGCGEQNMIRMAAPV 1005
 Qy 1023 TAVHVLDETEOMEXEFGLEKQGALELKKGYTQQLAPRQPSAFAFVRAEFTWLTAV 1082
 Db 1006 IATYLDLTTEQWETLGNRTEAVNQIVTGYAQMYKADHSHYAAFTNARSLLTAV 1065
 Qy 1093 VKVFLAVNLIA-IDSQVLGAVKWLILEKQKPDGVFQEDAPVIHQEMTGGERNNEKDM 1141
 Db 1066 VKVFAMAKRVAGISHIELICGVVRLILNRQPDGAFKFNAPVLSGTMOGGIQ-GAEEFV 1124

QY 1142 ALTAFLVLSIQEAKDICEQVNSLPGSITKAGDFLEANYNQLORSYTVTAIAGVALAQMOR 1201
Db 1145 YLTAFLVALLSSTKNDYNSLSSSKATNLLKXKYLQRPYTTALTAYALAAADQ 1184
QY 1202 LKGPLNKLFTAKDKNRWDFGKQLXNVEATSVALALLQLKDFDFVPPVWVWLNBOEY 1261
Db 1185 LNDRLVMAASTGRD--HWEEYNATHNIEGTSVALLAKMKKFDOTGPIVWLTQNF 1242
QY 1262 YGGVGSQTATFMYFOALAOYQKADPHQBLNDVLSOLPSRSSKTHRIHESASLLRS 1321
Db 1243 YGETYGGQTATFMYFOALAOYQKADPHQBLNDVLSOLPSRSSKTHRIHESASLLRS 1302
QY 1322 PETRENEGTVTAEKGOGTSLVVMYHAKAKDQL--TCNKFDLKVTKAPETEKRPQDA 1380
Db 1303 VETKLNQITVTASGDGKATMTILTFYNAQLQEKANVCNKFHLNVSVE--NIHLNAMGA 1359
QY 1381 KNTMLEICTRYRGDQATMILDSIMMTGFAPDQDQLKQANGVDVYISKYELDKAFSD 1440
Db 1360 KGALMKLCTRYLGEVDSMTIIDISLMTGLPDAEDLRLSKGVDRYISRYEVNMMQA 1419
QY 1441 RNTLIYLDKYSHSDCLAFKHQYFNVELIQGAVKVVAYYNLEESCTRYHPEKEDG 1500
Db 1420 KVAVIYLNKYSHSEDECLHFKLKHPEVGFIOGSKVYSYNLDEKCTKPYHDKGTG 1479
QY 1501 KLNKCRDLRCRCAEENCFOKSDDKVTLBERLDKACEPGVDVYKYLKVLQSLNDPDE 1560
Db 1480 LNKICIGNVCRCAGETCSSLNHGERIDVPLQEKACETNVDVYKYLKVLQSLNDPDE 1539
QY 1561 YMAIEQTKSGSDEVQVQQTFTSPIKCKREALKEEKHYLMWGLSDDFWGEKPNLSY 1620
Db 1540 YVMDVLEIKQGTDENPRAKTHQVISOQKQKQALNKLKNDYDLWGSRSDDLPTDKISY 1599
QY 1621 IIGKDTVWEHPEDEQDENQKQCDLGAFTESMVVRCGP 1662
Db 1600 IITNMTWIERPHEDECOEEBFQKLCDDFAQFSYLLTFEGCP 1641

RESULT 8

US-08-447-411-45
Sequence 45, Application US/08447411
Patent No. 5773243
GENERAL INFORMATION:
APPLICANT: FRITZINGER, DAVID C.
APPLICANT: BREDEHORST, REINHARD
APPLICANT: VOGEL, CARL-WILHELM
TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1755 S. Jefferson Davis Highway, Suite 400
City: Arlington
State: Virginia
Country: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,411
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/043,747
FILING DATE: 07-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5773243man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-101-0
TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248555 OPAT UR
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 1642 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-447-411-45

Query Match 48.9%; Score 4208; DB 1; Length 1642;
Best Local Similarity 49.5%; Pred. No. 0;
Matches 822; Conservative 322; Mismatches 482; Indels 36; Gaps 20;

QY 9 LLLLLLTHLPLALGSPWYSIITENILRLESEETMVLAEHDAQDQVPVTVTVHDPKGLV 68
Db 8 LVAALLIGPGSGHGLYLLITPAVLRTOETEQLVEAKGDSPTKQLDIFVHDFPRKQT 67
QY 69 LSSEKTVLTATNMGNVFTTIPANREFKSEKRNKFTVQATFGTV--VEKVLVLSQS 127
Db 68 LFQTRVDMNPAGQMLVTPTEIPA--KEVSTDSRONQYVVVQVT--GFQVLEKVVLSYQS 125
QY 128 GYLFIQTDKTIYTPGSTVLVRYLFTVNHKLLPVGRVTVMNIENPEGIPVKQDLSLSONQLG 187
Db 126 SFIFIQTDKGIYTPGSTVLVRYLFTVNHKLLPVGRVTVMNIENPEGIPVKQDLSLSONQLG 179
QY 188 VLPLSW--DPELVNMGQWKIRAYENSPQQVFSEFEVKEYVLSFSEFVIEPTKPYI 245
Db 180 -LNFPPFYNLPDLVSLGTWRIVAKYHSPEN--YTAFFDVVKYVLSFSEFVIEPTKPYI 237
QY 246 YNKGLEVTITAFPLGKVEGTAFFVIGIQDGEORISLPESLRIPEDDGSSEVVLSRK 305
Db 238 DGNENFHSITARYLYGEEVGVAFVLFVKIDDKKSIPDSLRIPEDDGSSEVVLSRK 297
QY 306 VLLDGQVNPRAEDLVCKSLVYSATVILHSGSDMWQAERSGIPVTSYQIHFTKPYK 365
Db 298 TFRS--RFPNLNLVGLTLYASVTVMTEGSDMWVTEQSGIHIVASPYQIHFTKPYK 355
QY 366 PGMPDLWVPTNPDGSPAVRVPVAVQGEDTVQSLTQSGDVAKLSINTHSPQKPSITVR 425
Db 356 PGMPYELTVYTNPDGSPAAHVPVSEAPHS--GTTLSGDTAKLILNPLNAQSPITVR 414
QY 426 TKKQELSEABQATRTMQALPYSTVGNNSNNYHLVRLTELREGETLNNVFNLLRMDRAHA 485
Db 415 TNGDLPREQAATKMTAAYQTQGGSGNYLHVATSTETKEGDNLPVKFQCEGCSL 474
QY 486 KIRYTVILMNKGLLKAGQVREPGQDLVPLSLTDDFIPSPRLVAYYTLIGAGORE 545
Db 475 QIKFYVILLINKGKIFKVGQPRDQONLVNWLHITPDLPISFRFVAYQV---GNNE 530
QY 546 VVADSVWVDKSCVGLSVVKGQSDROPVPGQOMTKIEGDHGAARVVLNAVDKGVFL 605
Db 531 IVADSVWVDKSCVGLSVVKGQSDROPVPGQOMTKIEGDHGAARVVLNAVDKGVFL 587
QY 606 NKXNKLTSQKIWDVVEKADIGCTPGSKDYGAGVSDAGLTFTSSSQCTQARAELOCP 665
Db 588 NDKYKISQAKIWDVVEKADIGCTPGSKDYGAGVSDAGLTFTSSSQCTQARAELOCP 647
QY 666 AARRR--SVOLTEKRMKVGY--PKELRKCCEBDGRENPMRPFSCORRTFISLGEACKV 723
Db 648 ANRRRSSVLLLDNSAKAAEFQDQDLRKCCEBDGRENPMRPFSCORRTFISLGEACKV 707
QY 724 FLDCCNVITELRROHARASHLGLARSLNLDIEDIAENIVSRSEFFESMLWVDEL--KEPP 782
Db 708 FLECCRYIKGVDRDQRESELEFJARDNEDGFIADSLISRSDFPKSLWLTKLTTEPN 767
QY 783 KNGISTKLNIPLKDSITTWELAVMSDKKICVADPFVTVMDFFIDRLRLPYSVVRN 842
Db 768 SQGISSTKMSFVLRDSTITVWVLAVSFTPTKGCVAEPYEVIRVMKVFVFFIDRLRLPYSVVRN 827
QY 843 EQVEIRAVLYNRQNOELKVRVLELHNPAFCSLATTKRHOCTITIPKSSLSVPPYVTP 902

Db 828 EQVEIRAILHNY-VNEDIYVRVELLYNPAFCASASTKQRYRQOQFPIKALSSRAVPFVIVP 886
QY 903 LKTGLOVEVEKAAVYHFFISDGVRSKLVVPPGIRNMKTAVRTLDPERLREGVQKEDI 962
Db 887 LEQGLHDEVEIKASVQKALSGVRSKLVVPPGIRNMKTAVRTLDPERLREGVQKEDI 945
QY 963 PPADLSQVDPDTESETRILLOTPVAQMTEDAVDAERLKHILVTPSGGEGQNMGTPTV 1022
Db 946 KAKLDRVPDTESETRILLOTPVAQMTEDAVDAERLKHILVTPSGGEGQNMGTPTV 1005
QY 1023 IAVHYLDETEQWKEFKLEKROGALELIKKGYYTQALFAPQSSAFAPVRAKAPSTWLTAYV 1082
Db 1006 IATYLLDTTEQWKEFKLEKROGALELIKKGYYTQALFAPQSSAFAPVRAKAPSTWLTAYV 1065
QY 1083 VKYPSLAVNLIA-IDSOVLGAGVWLLLEKQKEDGVFOEDAPVTHQEMIGGLRNNKDM 1141
Db 1066 VKYPSLAVNLIA-IDSOVLGAGVWLLLEKQKEDGVFOEDAPVTHQEMIGGLRNNKDM 1124
QY 1142 ALTAFLVLSIQEAKDCERQVNSLPGSIITKAGDFLEANNMLORSYTVIAIAGYALAQVR 1201
Db 1125 YLTAFLVLSIQEAKDCERQVNSLPGSIITKAGDFLEANNMLORSYTVIAIAGYALAQVR 1184
QY 1202 LKQPLINKFLTTAKDKRWEDPKQLVNVYATSYALLALLOKDFVPPVVRVNLQRY 1261
Db 1185 LNDRLVMAASTGRD--HWEENYNAHTNIEGTSYALLALLOKDFVPPVVRVNLQRY 1242
QY 1262 YGGYGSTQATFVFOALACYQKADPAPHOELNLDVSLQPSRSSKITHRIHWSASLLRS 1321
Db 1243 YGETYGTQATVMAFQALAEYEQMTHKDLNLDITIELDREVPVIRVINYENALLART 1302
QY 1322 BEYKENGFTVTAEGKQGLTSVWYMYHAKAKQDL-TGNKEDLVKTIKPAETETKRPQDA 1380
Db 1303 VETKLNODIIVTASGDKATMTILTFYNAQLQKANKVCKNFKELNVSVS---NIHLNANGA 1359
QY 1381 KNTMILICTRYRGDQDATMSILDI SMGTGFAPTDLDLQKLANGVDYRISKYELDKAFSD 1440
Db 1360 KGALMLKICTRYRGDQDATMSILDI SMGTGFAPTDLDLQKLANGVDYRISKYELDKAFSD 1419
QY 1441 RNTLIIYLDVSHSEDDCLAFKHQVYVVELLPQAGVYKAYVYNNLESCRTFYHPEKEDG 1500
Db 1420 KVAIIVYNNKSHSEDDCLAFKHQVYVVELLPQAGVYKAYVYNNLESCRTFYHPEKEDG 1479
QY 1501 KLNKLCRDELRCRAENCFIQSDDKVLTLEERLXACBPGVDYVYKTLVXQVLSNDFDE 1560
Db 1480 LLNKICIGVRCAGETCSLHQRIDVPLQIKERACETNVDYVYKTLVXQVLSNDFDE 1539
QY 1561 YVMAIEQTIKSGDEYVQVQORTFISPIKREALKLEKKHYLMKGLSSDFWGEKPNLSY 1620
Db 1540 YVMDVLEVIKQGHDKAPRATHQYISQRKQCALMLKNDYVYKTLVXQVLSNDFDE 1599
QY 1621 IIGKDTVWEHPDESCQDENQKQCDLIGAFTEGSMVWFGCP 1662
Db 1600 IITKNTWIERWHPDESCQDENQKQCDLIGAFTEGSMVWFGCP 1641

RESULT 9

US-08-662-227-35
; Sequence 35, Application US/08662227
; Patent No. 5922320
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; APPLICANT: BREDEHORST, REINHOLD
; APPLICANT: KOCK, MICHAEL
; APPLICANT: PRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVF
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA

ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/662,227
APPLICATION NUMBER: 530
FILING DATE: 14-JUN-1996
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-0107-OX
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 1648 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-662-227-35

Query Match 48.7%; Score 4194; DB 2; Length 1648;
Best Local Similarity 49.4%; Pred. No. 0;
Matches 821; Conservative 324; Mismatches 481; Indels 36; Gaps 20;
QY 9 LLLLLLTHLPALGSPVSIITENILFSEENMVLBAHQDGVPTVTVTHDFPCKLV 68
Db 8 LVAALLIGFPOSSGALYTLITPAVLTDEEQLVLAHQDSTPKQLDIFVDFPCKLT 67
QY 69 JSSEKTVLTATNMGVNTFTIPANREFKSEKGRNKPVTVAQTFGTQV-VEKVLVSLOS 127
Db 68 LQTRVDNPNAGGMLVPTTIEIPA-KEVSTDSQONQYVVQVT-GPQVLEKVLVSLOS 125
QY 128 GYLFQTDKTYTTPGSTVLYRIFTVNHKLLPVGRVTVMANIENPEGIPVAKQDLSLSONQUG 187
Db 126 SFLFQTDKTYTTPGSTVLYRIFTVNHKLLPVGRVTVMANIENPEGIPVAKQDLSLSONQUG 179
QY 188 VLPLSW--DIPELVNMGKIRAYENSPOQVSTEFVEKVLVSFEVIVTEKFFYI 245
Db 180 -LNFFWPNYLPDLVSLGTWRIYAKYHSPEN-YTAVFDVAKYVLPSEFVRLQSEKFFYI 237
QY 246 YNEKLEVTITATFELYGKVEGTAFVIFGIDGCEQISLPESLKRIPIEDGSGEVVLSRK 305
Db 238 DGNENFHSITATRIYIGEEVGVAFVIFGVKIDDAKSIEDSLTRPIIDGDKATLKR 297
QY 306 VLLDGVQNPRAEDLVGKSLVYSATVILHSGSDMVQAPRSOIPVTSPIQIHFKTKYFK 365
Db 298 TFRS--REFPNELVGHITLAVSVTMTESGDMVVTESGSIHIVASFYQIHFKTKYFK 355
QY 366 PCMPFDLAVFTNPDGSPAYRVPVAVQGEDTVQSLTQCGGVAKLSINTIPSOQPLSITVR 425
Db 356 PCMPFELVYVTPNPDGSPAAHVVPVSEAFHSY-GTLLSDGTAKLILNIPNAQSLPITVR 414
QY 426 TKKQSLSEAEQATRTMQALPYSVTGNSNNVILHLSVLTELRPGETLNVPLFLMDRAHEA 485
Db 415 TNHGDLPRERQATKSMATAIAYQTQGGSGNVLHVAITSTEIKPGDNLFPVNFVKNANSUK 474
QY 486 KIRYVYTLIMNKGRLKAGROVREPQDVLVPLISITTDPIFRFLVAVYVITLIGASQRE 545
Db 475 QIKYFTYLLNKGKIFKVGROPRDGGQNLVMTNLHITPDLIIPSFRVAYYQV---GNNE 530
QY 546 VTADSVYVDVDSQVGLSVLVKSGSEDRQPVPGQOMTLKEGHDGARVILVAVDKGVFL 605
Db 531 IVADSVYVDVDSQVGLSVLVKSGSEDRQPVPGQOMTLKEGHDGARVILVAVDKGVFL 587
QY 606 NKKNKLTQSKTWDVVEKADIGCTPGSGKDYAGVPSDAGLTFTSSSQOOTAQRAELQCPQ 665
Db 588 NDKYKISQAKIWDITIEKSDFCCTAGSGQNNILGVFEDAGLTFTSSSQOOTAQRAELQCPQ 647

```

QY 666 AARRRR-SVOLTEKMDVKYX-PKELRKCCEGRENHPRFSCQRTTRFISIGACKV 723
Db 648 AURRRSSVLLDSNASAAAFQDOLRKCCEGWHENPMGYTCERAKYIQGDACKAA 707
QY 724 FLDCNNYITELRQOHARSHLGLARSLNDEIIAENIVSRSEPPESWLMNVDEL-KEPP 782
Db 708 FLECCRYIKGVADENQRESELFLARDNEDGFIADSDIISRSDFPKWMLTKDLTEPN 767
QY 783 KNGISTKLMNFKLDSITTWELAVSMKDKGICVADPFVYVMDFFIDRLRPLYSVRN 842
Db 768 SGISSKMSFYLRSITTWVLAWSFTPTKGCIVABPFYIRVMKVFIFDLQNPYSVKN 827
QY 843 EQVEIRAVLYNRQONELKVRVELLHNPAPFCSLATTKRHQQITITPPKSSLSVPYVIVP 902
Db 828 EQVEIRAILHNY-VNEDIYVRVELLYNFAFCSASTKGQRYRQOPFIKALSSRAVPFVIVP 886
QY 903 LKATGOEYEVKAAYVHHFISGVRSKLVKVPGEIRNMKNVAVRTLDPERLGRGVQKEDI 962
Db 887 LEQGLHDEIIRKASVQEAALWSGVKRLKVPGEVQ-KSIVTVKLDPRAGVGGTQLEVI 945
QY 963 PPADLSQDQVPTSETRILLQSTPVAQMTEDAVDAERLKHLYVTPSCGCEQNMIGMTPTV 1022
Db 946 KARKLDRVPTDEIETKIIQGDPAQIENSIDGSKNLHLIITPSCGCEQNMIRMAAPV 1005
QY 1023 IAVHYLDETEQWEKFGLEKRGALIKKGYTOQLAPQSSAFAPKVPKAPSTWLTATV 1082
Db 1006 IATYVLDTEQWETLGINRTEAVNQIVTGAQMVYKKAADHSAFTNTRASSWLTATV 1065
QY 1083 KVVFSLAVNLJA-IDSOVLGAVKWLILEKQKPGVQEDAPVHOEMIGGLRNNNEKDM 1141
Db 1066 KVVFANAAKAVAGISHEIICGVRLWILNRQOPDGAFAKNAFLVSGTMQGGIQ-GABEEV 1124
QY 1142 ALTAFLVLSLOAKDICEQVNSLPGSITKAGDFLEANYVNLORSYTVATAGYALQAMGR 1201
Db 1125 YLTAFILVALLSKTICNDYVNSDSSIKKATNLKLYEKLQRPYTTATATALAAAOQ 1184
QY 1202 LKGPLNLKFLTTAKDKRWEDPGKQLVNVETASYALLALQLKDFDFVPPVVRWLNQRY 1261
Db 1185 LNDRLVMAASTGRD--HWESYNATHNIEGTSYALLALLMKKFKDQGTGIVSWLTDQNF 1242
QY 1262 YGGVGSSTCATPMVFOALAQKADPOHOLNLDVSLQLPSRSKITHRHWSASLLRS 1321
Db 1243 YGETGQIQATVNAFAQAEYEQIOMPHKDLNLDITIELDPREVPIRYIRINENALLAT 1302
QY 1322 BETKNEGFTVABGKGGTSLVVMYHAKAKDOL-TCNKFPLKVTIKKAPETERKPDQA 1380
Db 1303 VETKLNQDITVTASGDGKATILTFYNAOLQEKANVCNFKLNVSE--NIHLNAMGA 1359
QY 1381 KNTMLEICTRYRGDQDATTMILDI SMWTGPAPDQDLDKOLANGVDRIYSKYELDKAFSD 1440
Db 1360 KQALMKIKCTRYLGEVDSMTIILDI SMLTGFLPAEDLTLKSGVDRIYSRYEVDNNMAQ 1419
QY 1441 RNTLIYLDKVSHEDDCLAPKXHQYNVELIQGAVKYAYYNLEESCTRFYHPEKEDG 1500
Db 1420 KVAVIYLNKVSHEDECLHPKILKHFEVGFIPQGSVKVSYYNLDEKCTKFVHPDKGTG 1479
QY 1501 KNLKLCRDLRCBAENCFIQKSDDKYTLLEERLDKACEPGVDVYVTKLVKQLSDNDFE 1560
Db 1480 LNKLCIGNVRCACETCSSLNHQERIDVPLQIEKACETNVDVYVTKLRLBEEQDNDI 1539
QY 1561 YMAJEQITKSGSEVQVQORTTISPICKREALKLEBKHYLWGLSDDFMGEKENLSY 1620
Db 1540 YVMDVLEVIKQTDENPRAKTHQVYSQRKQOEALNKNVNDYLLINGSRDILLPTKDKISY 1599
QY 1621 IIGKDTWVEHPDEDECOENOKOCODLGAFTESMVVFGCP 1662
Db 1600 IITKNTWTERPHDEDECOEEFQKLCDDPAQFSYTLTEFGCP 1641

```

RESULT 10

US-09-017-947-35

; Sequence 35, Application US/09017947

```

; Patent No. 6303754
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; APPLICANT: BREDEHORST, REINHORST
; APPLICANT: KOCK, MICHAEL
; APPLICANT: FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVF
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/017,947
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/662,227
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-0107-0X
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1648 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-017-947-35

```

Query Match 48.7%; Score 4194; DB 3; Length 1648;

Best Local Similarity 49.4%; Pred. No. 0;

Matches 821; Conservative 324; Mismatches 481; Indels 36; Gaps 20;

```

QY 9 LLLLLLTHPLALGSPMYSIIITENILRLSEETMVLREHDAQDVPVTVVTHDFPKKLV 68
Db 8 LVAALLTGFPGSHGALYLTITPAVLDTDTESQILVEAHGSDTPKQLDIFVHDFPKKLT 67
QY 69 LSSEKTVLTPTAHMGMVNTFTIPANREPSEKGRNKFVTVQATFGTV-VKQVVLVSQS 127
Db 68 LFQTRVDMNPAGMLVTPTEIPA-KEVSTDSRQNVVVQVT-GPQVLEKVVLLSYQS 125
QY 128 GYLFIQDQKTYTPGSGVLYRIETVNHKLLPVCRVTVMNIENPEGIPVQKDSLSQNLG 187
Db 126 SFLFIQDQKTYTPGSGVLYRVFSDHHTSKMKTIVIVFQTEGPEGLVSSNSVD----- 179
QY 188 VLPLSW--DIPELVNMGMKIRAYENSPQVSTFEFVKEYVLPSEFVIEPTKPYI 245
Db 180 -LNFPPYNDLPDLVSLGTWRIVAKYHSPEN-YTAYFDVRKVLPSFEVRLQSEKFFYI 237
QY 246 YNEKGLVTLTARFLYKKGVEGTAFVFIQDGEORISLPESLKIPIEDSGSEVVLGRK 305
Db 238 DGNENFVHSITARYLYGEEVGVAFVLFQVKKIDDAKKSIPDSTLRIPIDGDKATLRD 297
QY 306 VLLDGVQNPRAEDLVGKSLYVSATVILHSGSDMVQAERSGIPVTSYQIHFTKTPKYPK 365
Db 298 TFRS--RFPNLLNELVGHITLYASVTVMTEGSDMVVTEQSGIHIVASPYQHFTKTPKYPK 355
QY 366 PGMPFDLMVFTVNPFDGSPAVRVPAVQGEDTVQSLTQSGDVAKLSINTHPSQKPLSITVR 425

```

Db 356 PGMPELTVTVTDPGSPAARFVWSEAFHSM-GTTLSDGTAKLILNIPLNAQSLPIVTR 414
QY 426 TKQELSEASQARTWCAALPYSVTGNSNYLHLSVLRLTELRLGETLNVNLEMLDRMRAHEA 485
Db 415 TNGGDLPREQATKSTATAIYQTGGSGNYLHVATSTBEIKPGOMLPNPNVKNVANSXLK 474
QY 486 KIRYTYILMKGELLKAGQVREPOODLVPLSLTTFIPSPFLVAYYTLIGASQORE 545
Db 475 QIKYFTYILNKGKIFKVGQRPRDQGNLVTNLNLHTPLIPSPRFVAYVQV---GNNE 530
QY 546 VVADSVMDVYKSCVGLVVKSQSDEDRQPVQSQWTLKIEGDHGARVLVAVDKGVFVL 605
Db 531 IVADSVMDVYKSCVGLVVK---GDLIOMFGAMKIKLEGDFGARVLVAVDKGVFVL 587
QY 606 NKXNKJTSKIDVVPKADIGCTPGSGDYAGVFSAGLTFTSSSGQQTARAELOQCPQ 665
Db 588 NDKYKISQAKIWDITIEKSDPGTAGSGQNNLGVFSDAGLALTSTNLNTKORSAAKCPQ 647
QY 666 AARERR-SVOLTEKMDKVGKY-PKELRKCCDGMENPMRPSCORRTRFTISLGEACKV 723
Db 648 ANRRRSVLLLSNASKAABFQODLRKCCEDVHENPMGYTCERAKYIQSGACKAA 707
QY 724 FLDCNVTIELRQHAARASHGLIARSNLDEDIIAENINVSREBPESMLNVEDL-KEPP 782
Db 708 FLECCRYIKGVDRDENORESELFARDNEDGFIADSDIISRSDFPKWMLTKOLTBEPN 767
QY 783 KNGISTKLMNIFLKDSTITWEILAVSMDSKIGICVADPEVTVWODFIDLRLPYSVVN 842
Db 768 SQGLSSKTMFYLRDSITWVLAIVFTFTKIGICVABPYEIRVMKVFFIDICMPYSVKN 827
QY 843 EQVEIRAVLYNQNOELKYRVELLNPAFCSLATYRRHQQTITIPKSSLSVPPYIVP 902
Db 828 EQVEIRAILHNY-VNEDIYVYRVELLYNPAFCSTASTKQRYQROPPIKALSSRAVPFVIVP 886
QY 903 LKTCLOVEVKAAYVHHFISDGVKSLKVUPEGIRNKTAVVRLDPERLREGVQKEDI 962
Db 887 LEQGLHVDVEIKASQVBAISDGVKSLKVUPEGVQ-KSIVTIVKLDPRAGVGGTQLEVI 945
QY 963 PPADLSQVPTSETRILLQGTTPVQMTEDAVDAERLKHILVTPSGGGEONMGWPTV 1022
Db 946 KARKLDDRVPTDIETKIITQSGPVAQIENSIDGSKNLHLITPSGGGEONMRMAAPV 1005
QY 1023 IAVHLDTEBQWERFGEKQGALELILKGYTQOLAFROPSSAFKAFVKRAPSTWLTAYV 1082
Db 1006 IATYLDTTQWETLGINRTEAVNQIVTGYAQWYKKAHSHYAAFTNRASSSWLTAYV 1065
QY 1083 VKPFLAVNLIA-IDSOVLGAVKWLILEKQKPDGVFQEDAPVHCEWIGGLRNNEKDM 1141
Db 1066 VKVFAMAAKVAGISHSIILGGVRLILNRQPDGAFKENAPVLSGTWGGIQ-GABEEV 1124
QY 1142 ALTAFLVLSIQEAKDICEOVNSLPGSITKAGDFLEANYMNLQSYTVIAIAGYALAQMR 1201
Db 1125 YLTAFLVLSLEKTCNDVNSLDSIKKATNLLKYEKLQRPYTTALTAYALAAADQ 1184
QY 1202 LKGPLLNKFLTTAKKNRWDPKQKLYNVEATSYALLALLQLKDFDPVPPVVRWLNBOY 1261
Db 1185 LNDRLVMAASTGRD--HWBEYNAHTNIEGT-SYALLALLKMKKFDQGPVVRWLTQNF 1242
QY 1262 YGGYGTQATVMAFAQAEYQIOMPHKDLNLDITIELPDREVPVRYRYENYENALLART 1321
Db 1243 YGETYGTQATVMAFAQAEYQIOMPHKDLNLDITIELPDREVPVRYRYENYENALLART 1302
QY 1322 EETKENEGFTVTAEGKQGLTSVVWYHAKADQL-TCNKFDLKVITKPAETEKRPQDA 1380
Db 1303 VETKLNQDITVTASGDGKATMTILTFYNAQLOEKANVKNKFLNVSVE---NIHLNANGA 1359
QY 1381 KNTMILICTRYRQDQATMSIILDSMTGFAFPTDLDLQKLANGVDRYISKYELDKAFSD 1440
Db 1360 KGALMLKICTRYLGEVDSTWITIDISVLGTGLPDAEDLTELKSGVDYRYSRYEVDNNMAQ 1419
QY 1441 RNTLIIILDKVSHSDDCLAFKVOHYENVTELIOPGAVKYAYNLEESTRYHHEKEDG 1500

1420 KVAVIYINKVSHSEDECLHPKILKHPFVGFIOQPSVKVYSYNNLDEKCTKFHPDKGTG 1479
1501 KLKLCRDELRCRAEENCFTQKSDDKVTLEERLDKACEPGVDYVYKTRLVKQLSNDPDE 1560
1480 LNKICIGNVCRCAGETCSSLNHQERIDVPLQIEKACETNVDYVYKTRLVKQLSNDPDE 1539
1561 YMALEQITIKSGSDEVQVQORTFISPIKCBALKLEEKHYLMWGLSGDFWGEKENLSV 1620
1540 YMDVLEVIKGTDENPRAKTHQVYSORKCOEALNKVNDYLYLWGSRLDLPKDKISY 1599
1621 ITGKDTWVHEPDEDEQDENQKQCDLGAFTESMVVFGCP 1662
1600 IITKXTWIERPHEDEQEBEFQKLCDDFAQPSVLTTEFGCP 1641

RESULT 11
US-09-925-442-35
; Sequence 35, Application US/09925442
; Patent No. 6607897
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; KOCK, MICHAEL
; BSEDEHORST, REINHORST
; FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVP
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OSLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/925,442
; FILING DATE: 10-AUG-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/017,947
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-0107-0X
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1648 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-925-442-35

Query Match 48.7%; Score 4194; DB 4; Length 1648;
Best Local Similarity 49.4%; Pred. No. 0;
Matches 821; Conservative 324; Mismatches 481; Indels 36; Gaps 20;

QY 9 LLLLLTHLPALGSPMYSIITPNILRESEETWLESAHQDQGVPTVTVHDPGKGLV 68
Db 8 LVAALLGPPGSHGALYTLITPAVLRDTCTBEQILVEAHGDSPTKQLDIFVHDFPRKQKT 67
QY 69 LASEXTVLTPATHNMGVTFITPANREFKSEKNGKVTYQATFGTQV-VEKVVVLSLQS 127

68 LPQTRVDMNPAAGMLVTPTIEIPA-KEVSTDSRQNVVVQVVT-GPOVLEKVVLLSYQS 125
 128 GYLFIOTDKTYTTSSTVLRIRIFVHKLIPVGRVMTNVALENPEGIPVQDLSLSONQLG 187
 126 SFLFIQTDKGIYTPGSPVLRVFSMDHHTSMKNTVIVFQTEGFLVSSNVD----- 179
 188 VLPLSW--DIPELVMQWIRAYENSPOQVSTEFEBEYVLPSEFVIVEFTKPYI 245
 180 -LNFEPYNLPDLVSLGTWRIVAKYHSPEN-VTAYFDRKVVLPSEFVRLQPSKFFYI 237
 246 YNEKGLVTTIARLYGKVEGTAFVIFGLODGEQIRISLPESLKRIPIEDSGSEVVLGRK 305
 238 DGENFHVSIITARYLYGEEVGFVLFVGFVIDAKKSPDSITRPIIDGCKATLKR 297
 306 VLLDGVQNPRAEDLVGKSLYVSATVILHSGSDVMQVABERGIPVITSYQIHTKTPKPK 365
 298 TFRS--RFPNNEVLGHITLASVTMTESGSDVMVTEQSGIHIVASPYQIHTKTPKPK 355
 366 POMPFDLMVFTNPDGSPAVRVPVAVOGEDTVQSLTQDGVAKLSINTHPSOKPSITVR 425
 366 POMPVELIVYTNPDGSPAAHVFPVVSAPHSM-GTILSDGTAKLILINIPNAQSPITVR 414
 426 TKKQELSEABQATRMQALPSTVGNNSNHLISVLRTELPGETLNVNPLLRMDRAHEA 485
 415 TNHGDLPREARQATKSMATIAIYQTOGSGNYLHVAITSTEIKPGDNLFPVNFVNGVANSILK 474
 486 KIRYVYLLIMKGLLKAGROVREPQDVLVPLSIITDFIPSPFRVAVYITLIGASGORE 545
 475 QIKFYIILNKGLFKVGRPRDQGNLVTMNLHITPDLIPSPRVAYVQV---GNNE 530
 546 VVADSVWVDVDSVGLSVKSGOSEDROPVPGOMTLKIEGHGAVRVVLVAVDGVFVL 605
 531 IVADSVWVDVDTGNTGLVWV--GDNLIQMPGAAMKIKLEGDFGARVLVAVDCAVYL 587
 606 NKKKLTOSKIWDVVVEKIDICTPGSGKDYAGVFSOAGLTFTSSSQOQTAQAEALQCPQP 665
 588 NDKYKISOAKIWDITKSDFTAGTAGSQNNLGHFEDAGLALTSTNLTKQSAKCPQP 647
 666 AARRR-SVQITKRMKVKY-PKELKCCEDGNRENPMRFSQORTRFISLGEACKV 723
 648 ANRRRSVLLDSSNAKAAEFQODLRKCCEDVMHNPNGYTCERAKYIOEGDACKAA 707
 724 FLDCNRYTELRRQARASHLGLARSLNIDELIAENIVSRSEFFPSWLNWVEDL-KEPP 782
 708 FLECCRYIKGVDRNQRESELFLARDNEDGFIADSDIISRSDFPKWMLWLTOLTEEN 767
 783 KNGISTKLMLFLKDSITTWELILAVMSDKKICVADPPVTVMQDFFIDLPLPYVYN 842
 769 SQGISKMSFVLRSDIITWVLAVSFTPTKIGICVAPPEYRVMKVFFIDLQMPYSVWKN 827
 843 EQVEIRAVLYNRQNELKVRVLLHNPAPCSLATTKRHQQTITIPPKSSLSVPYVIP 902
 828 EQVEIRAILNY-VNEDIYVRVELLYNPAPCSASTKGORYRQPPKALSSRAVPFVIP 886
 903 LKTLGOEVEKAAVYHHSIDGVKSLKVPYEGTRMKNKYAVVTLQPERLGRGCKEDI 962
 887 LEQGLHDEIRASQVQALWSGVRRKKLVVPEGVQ-KSIVTVLKLDPRAKGVGTGLEVI 945
 963 PPADLSQDVPDTSESTRILLQSTPVAQNTEDAVDAERLKHIVTPSGCGEQNMIGMPTV 1022
 946 KARKLDRVPDTETETIKIIQGDVPAQIENSIDGKLNHLIITPSGCGEQNMIRMAAFV 1005
 1023 IAVHYLDEQWEKEGLEGKROGALILIKGYTQOLAFQPSAFAFVKEASTWLTAV 1082
 1006 IATYVLDTEQWETLGINRTEAVNQIVTGAQVYKKAACHSAFTNRASSWLTAV 1065
 1083 VKVFLSALNLTIA-IDSQVLCGAVKWLILEKQKPGVFOEDAPVHQMETGGJRNNEKDM 1141
 1066 VKVFAAMAKWAGISHETICGVRWLILNRQOPDGAFAKFNAPVLSGTWQGGIQ-GAEEV 1124
 1142 ALTAFLVLSQKXIDCEOVNSLPGSITKAGDPLEANYVNLORSYTVLAVYALQMR 1201
 1125 YLTAFILVALLKSTICNDYVNSLDSKIKKATNVLKYEKLQRPYTTALTAYALANAO 1184

1202 LKGPLLNKFLTTAKKNWEDFGKOLYNNVEATSVALLALLQLQDFDPVPPVWRWLNQRY 1261
 1185 LNDRLVMAASGRD--HWEYNATHNIEGTSTVALLALLKMKKFDOTGPIVRWLTQNF 1242
 1262 YGGYSGSTQATPMVFOALAQYQKADPDHQBELNLDVLSQLPSSSKITHRIHWESASLLRS 1321
 1243 YGTYGOTQATWAFQALAEYEQIOMPHKDLNLDITIELDPREVPPIRYR-NYENALLART 1302
 1322 EETKEGEGFTVABGGGOTLSVWVMYHAKAQGL-TCKNFKDLKVTIKPAPETEKRPQDA 1380
 1303 VETKLNDQITVTSAGDGRATWILTFYNAQIQEANCNKHFLNVSVE--NIHLNAMGA 1359
 1381 KNTMILEICTRVRGQODATMSILDSMTGFPAPTDLLKQLANGVDYRISKYELDKAFSD 1440
 1360 KGALMKIKITRYLGEVDSMTIID:SMLTGFLPDAEDLTRLSKGVDRYISRYEVNNMAQ 1419
 1441 RNTLIYLDKVSJSHSDCLAFKVOYFNVELIQCAVKVAYVYNLESCFRFYHPEKEDG 1500
 1420 KVAVIILANKVSHSEDECLHPKILKHFEVGFQPGSVKVSYYNLDKCTKFVHPDKGTG 1479
 1501 KLKLCRDBLCBACENCFIQKSDDKVTLERLDKACEFGVDYVYKTLVKVQLSNDPDE 1560
 1480 LMKICIGNVRCAGETCSSLNHQBRIIDVPIQIEKACETNVVYKTKLURIEQCGNDI 1539
 1561 YIMAIRQTTIKSGSDEVQVQQRFTFSP:KCREALKLEBKHYLMWGLSSDFWGEKPNLSY 1620
 1540 YVMDVLEVIKQGTDENPRAKTHOYISQKQCEALNKLKVNDDYLIWGSRLDLPKDKISY 1599
 1621 IIGKDTWHEWBEDECODEENKOCODLGAFETSSWVVEGCP 1662
 1600 IITNTWIERWBEDECODEEFCQKUCDDFAQFSTLTBEFGCP 1641

RESULT 12
 US-08-447-411-76
 ; Sequence 76, Application US/08447411
 ; Patent No. 5773243
 ; GENERAL INFORMATION:
 ; APPLICANT: FRIZINGER, DAVID C.
 ; APPLICANT: BREDEHORST, REINHARD
 ; APPLICANT: VOGEL, CARL-WILHELM
 ; TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
 ; NUMBER OF SEQUENCES: 81
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
 ; ADDRESSEE: P.C. S. Jefferson Davis Highway, Suite 400
 ; STREET: 1755 S. Arlington
 ; CITY: Arlington
 ; STATE: Virginia
 ; COUNTRY: U.S.A.
 ; ZIP: 22202
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/043,747
 ; FILING DATE: 07-APR-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Oblon, No. 5773243man F.
 ; REGISTRATION NUMBER: 24,618
 ; REFERENCE/DOCKET NUMBER: 1126-101-0
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 413-3000
 ; TELEFAX: (703) 413-2220
 ; TELEX: 248855 OPAT UR
 ; INFORMATION FOR SEQ ID NO: 76:

SEQUENCE CHARACTERISTICS:
LENGTH: 1333 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-447-411-76

Query Match 41.3% Score 3559; DB 1; Length 1333;
Best Local Similarity 51.3%; Pred. No. 0;
Matches 683; Conservative 255; Mismatches 372; Indels 22; Gaps 13;

QY 334 SGGDWAOBESGPIVTPSPQIHFTKTPKYPKPGMPFDMVFTNPDGSPAYRPV---A 390
DB 3 SGGDWMTQESGHIHTSPQIYFTKTPKYPKPGMPFELTVVVTKPDGSPAHPVYVSEA 62
QY 391 VQGEDTVQSLTQDGGVAKLSINTHPKXPLSTVTRKQELSEAEQATRTWQALPYSTVG 450
DB 63 IHSECT---TLSDGTAKLFLNTPQNAQSLPTVTRNHGDLPREQAISKMTATAYQTQ 118
QY 451 NSNYHLVRLTRELPGETLNVNFLRMDRAHEAKIRIYTYLNMKRLKAGQVREP 510
DB 119 GSGNYLHVATITETKPGDNLVFNVRGNANSLNQIKYFTYILNKGKIFKVRGRHRG 178
QY 511 GQDLVLPISITTFPSRLVAYTYLIGASQREVWADSVMDVKDSCVGLVVKSGQS 570
DB 179 GENLVNWLHITPDLPSPFAVYQV---GNRIVADSVMDVKDTCMGTLVVKGATS 234
QY 571 ED-RQVPVGGQMTLKIEGHGARGVAVLVAVDKGVFNKKNKULTQSKIMDVVEKADIGCTP 629
DB 235 RDNRIQMGAAKIKLEGPGAWIGLVAVDKRABYVINDKIKTSQAKIWDITIEKSDFGCTA 294
QY 630 GSGKXDYAGVSDAGLFTSSGQOTQRAELQCPQAPARRRR-SVOLTEKMDVKGY-P 687
DB 295 GSGQNNLVGFEDAGLATTSTNLNTKQSAKCPQAPARRRRSSVLLDLSNASTAAQFD 354
QY 688 KELKCCEDGMENRFRFCORTRFISLGEACKVFLDCNNTYELRRQARASHGLA 747
DB 355 QDLKCCEDGMENRPHGHTCEKREKXIQEGDACKAFLPCCCHYKGIQDGNKRESEFLA 414
QY 748 RSNLDEDIABENIVSRSPFSSWLNWVDELKEPKN-GISTKLMNIFLKDSTITWELLA 806
DB 415 RSDFEDDLFCGNIITSRSPFSSWLNWVDELKEPKN-GISTKLMNIFLKDSTITWELLA 474
QY 807 VNSDKKGIQVAPFEVWMDFFIDRLPYSVNRNEQVEIRAVLYNRQNELKYVEL 866
DB 475 VGLSPTKGIQVAPFEVWMDFFIDRLPYSVNRNEQVEIRAVLYNRQNELKYVEL 533
QY 867 LHNPAFCSLATTERRHQCTITIPPKSSLSVPYVIVPLKTLQEVVEVKAAYVHHFISDVR 926
DB 534 LYSAPFASSTESQVRQELPIKALSSRAVSPVIVPLSQGLHDVEVTASVQGLMSDGVK 593
QY 927 KSLKVPSEGRIMKNTAVATLDERLGRGVQKEDI PPADLSDOYDPTESRTRILLGTP 986
DB 594 KKLKVPSEG-EWKSIVTIELDHTKIGTQVLELVKANKLNDVRPDEITETIKTIQDGP 652
QY 987 VAQMTDAVDAERKLHIVTPSCGQNNMIGMTPPIAVHYLDETEQWKEFGLKROCAL 1046
DB 653 VAQTIENSIDGSKLNHLIITPFCGQNNMIRMTAPVIATYVLTDTQQWETLGINRTEAV 712
QY 1047 ELIKKGYTQOLAFROPSSAFAPVKEAPSTWLTAVVAVSLAVNLIA-IDSOVLCAVVK 1105
DB 713 NQMTGVAQLVKKADHSAFETNSASSSWLTAVVKIPALAAKIVKDIKHEIVCGMR 772
QY 1106 WLILEKQKPGVQEDAPVTHQBMIGGLRNNEKDMALTAFLVLSLOAKDICBEQVNSL 1165
DB 773 WLILNRQRTDGVFRENAPVLFQTMQGIQ-GAEPGSLTAFILVALLSRSICNAYNIL 831
QY 1166 PGSTTKAGDPLEANYNLRQSVYVAVGALAAQMKLKGPLLNKFLTTAKDNWEDPGK 1225
DB 832 DSSISKATDYLKKYKELQRPYITALTAYALAAAEELNDRV-LMAASTGGRNWEFPA 899
QY 1226 QLYNVEATSVALLALLQLQDFDVPVFWLNEQRYGGYGGTQATFWQALAQYOKD 1285

DB 890 HTHNIEGTSVALLALLKMKKFEAGPVQVWLIDQQYGGTYGQTQATVWMPQALAEYBIQ 949
QY 1286 APDQELNLDVSLQPSRSKITHRIHWSASILLSEETKENEGETVTAEGQQTLSVW 1345
DB 950 MPTHKDLNLDITIELPDREVPFIRYINYNALIAQTVETKLNEDFTVSASGDKATMIL 1009
QY 1346 TWYHAKAK-DQLTCNKFDLAVTIKPAPETEKRPQDAKNTMILEICTRYRGDODATMSILD 1404
DB 1010 TVYNAQLREDAVNCNPHLDVSVENQLNLKAKGAKGALKIKICTRYLGEVDSVTWIID 1069
QY 1405 ISMTGTAPDTRDLKQLANGVDRIYISKYELDKAFSDRNLIIYLVKSHSEDDCLAFKVH 1464
DB 1070 VSMLTGVPDTELDLRELKSGVDRIYISMPFINNNMAQKGTVIIYLDKXSHSEDECLHFKIL 1129
QY 1465 QYFNVETIQGAVKAVVAYNLEESCTRFYHPEKEDGKLNKCRDBLCRCAEENCFCIQKSD 1524
DB 1130 KHFEVGIQPGSVKYVSYNLDKCTKIYHPDEATGLANKICVGNVCRCAETCSLLNQ 1189
QY 1525 DKVTLERLQKACBPVDVYVYKTRLVKQLSNDPDEYIMAEIQTIKSGSDEVOVQOQRTF 1584
DB 1190 KNTVTRQRIQKAFDPNDVYVYKTLRIEKGNDIYVMDVLEVLKQSTQDQNCQVXVQY 1249
QY 1585 ISPTKCEALKLEKHYLMWGLSSDPWGEKKNLSYIIIGKDTWVSHWPEDEECQDEENQK 1644
DB 1250 VSQKCEALNLVNDYLLWGPSSDLMPKDKISYLIITNTWIERWPHEDKQDEEFQK 1309
QY 1645 QCQDLGAFTESM 1656
DB 1310 LCDDFALFSYAM 1321

RESULT 13
US-08-662-227-34
Sequence 34; Application US/08662227
Patent No 5922320
GENERAL INFORMATION:
APPLICANT: VOGEL, CARL-WILHELM
APPLICANT: BREDEHORST, REINHORST
APPLICANT: KOCK, MICHAEL
APPLICANT: FRITZINGER, DAVID
TITLE OF INVENTION: RECOMBINANT PROCVF
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/662,227
FILING DATE: 14-JUN-1996
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-0107-0X
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 1333 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

Query Match 41.1%; Score 3536; DB 3; Length 1333;
Best Local Similarity 51.2%; Pred. No. 0;
Matches 683; Conservative 254; Mismatches 372; Indels 24; Gaps 15;

QY 334 SGGDMVAERSGIPVTPSPQIHFTKTPKYPKPGMPFDMVFNTPDGPSPAYRPV--A 390
DB 3 SGGDMVAERSGIHVTPSPQIHFTKTPKYPKPGMPFDMVFNTPDGPSPAYRPVSEA 62
QY 391 VQSDTVOISUTQDGVAKLSINHPKOKPLSTIVTKKQELSEABOARTWOALPYSTVG 450
DB 63 IHSEGT---TLDGAKLFLNTPONAQLPIVTRNHGDLPRERQAISKSTAYATQ 118
QY 451 NSNYHLVRLTELAPGETLNVNFLRMDRAHEAKIRYIYVLMNKGLLLKAGQVREP 510
DB 119 GSGNYLVAITSTEIKPGDNLPNFNVGNANSLNQIKYFTYIILNKGIKFKVGRHRGD 178
QY 511 GQDLVLPSTITTFPSFRLVAYYTLIGASQREV-VADSVVVDVKSCVGSLLVWKSQ 569
DB 179 G-NLVTNHLHTLIPSRFVAYQV---GNNEIEVADSVVVDVKDTCGTLVVKAT 233
QY 570 SED-RQVPQOQMTLKIEGDHGRVVLVAVDKGVFLNKKKLTQSKLWDVWVVKADTCT 628
DB 234 SRDNRIQMPGAAMKILKEDGPGAWIGLVADVAEYVLDNKKYISQAKIWDITIEKDFGCT 293
QY 629 PGSGDYAGVFS DAGLTFITSSSQQAQAEIQCQPAARRRR-SVQLTEKRMKGVKY- 686
DB 294 AGSQNNLGVFEDAGLALITSTNLTKQSAKACQPAARRRRSSVLLDLSNASKAAQFQ 353
QY 687 PKELKCCBGMENPMRSCQARTFISLGACKVFLDCNVIYTELBRHARASHLGL 746
DB 354 DQLRKCCBGMENPMHGTCEKERYIQEGDACAAFLCCHYIKGIQDDNKSESLFL 413
QY 747 ARSLDEDIIAENINYSRSEFPEMLNVEDLKEPPKN-GISTKLMTFLKDSITTWEIL 805
DB 414 ARSDFEDDLFGEGNITSRSDFPESLMLMEQLSEHSPKSGISKIYVPYLRDSITTWELL 473
QY 806 AVNSDKKGCIVADPEVTVNQDFIDLPLYSVYVNEQVEIRAVLYNRYONQELKVRVE 865
DB 474 AVGLSPTRKICVAPETVEITVWMDFFIDLPLYSVYVNEQVIRAVLYNRY-ADKDIYVRVE 532
QY 866 LLHNPFCSLATTKRHOQTITIPPKSSLSVPYIVPLKTLQIEVEKAAVYHFFISDGV 925
DB 533 LLYSPAFCSASTESQRYREOLPIKALSSRAVSFVIVPLEOGLHDVEVTASVOGELMGDGV 592
QY 926 RKSILVYVEGIRKNTKAVRTLDPERLREGVQKEDIPPADLSDOVPTSETRILLQGT 985
DB 593 KKKLVYFEG-EWKSIVTIELDPHTKIGCGTQVSLVANKLNDRPDTEIETKITIQGD 651
QY 986 PVAQMTEDAVDAELKHLIVTPSGCEQNMIGMTPTVIAVYLDTEQWEXFLEKROGA 1045
DB 652 PVAQTIENSIDGSKLNLHIIITPFGCEQNMIRMTAPVIATYVLDTQOWETLGNRTEA 711
QY 1046 LELIKGYTQOLAFROPSSAFVVRAPSTWLTAVVVFSLAVNLIA-IDSOVLGAV 1104
DB 712 VQJMTGYAQLVYKXADHSYAFTNSASSNLTYVVKIFALAKIVKNDINHIVCGM 771
QY 1105 KWILEKQKPDGFQEDAPVIHQMIGGRNNKDMALTAFLVLSIOAKDICEOVNS 1164
DB 772 RWLLNQRDQDGFRENAPVLFOTMQGIGQ-GAEPEGLTAPILVALLESRSICNAVINI 830
QY 1165 LPSGITKAGDFLEANTYMNQORSYTVAIAGYALQAGMLKPLNKLFTTAKDKRWEDPG 1224
DB 831 LDSSISKATDYLKKYKELQRPYTTALTAYALAAERLNDORV--LMAASTGRNRWEPN 888
QY 1225 KOLYNVETAYALLALLOLKDPEFVPRVNLNBOYRGYGGVGSTOATFWVFOALQYK 1284
DB 889 ARTHNIEGSTYALLALLOLKKFVEAGPVVQVLDIQGYGTYGQYQATVMVFOALAEYI 948
QY 1285 DAPDQELNDVSLQFSSSKITHRIHWESASLSRSEBETKENEGFTVTAEGKQOGLTS 1344
DB 949 QMPHDKLNDLITIELPDRVPRVRYRINYNALLAQVETKLNEDFTVVSAGDGKATWTI 1008
QY 1345 VTMWAKAK-DQLTCNKFDLKVTKPATEKRPQDANKTMLEICTRYRQDQATWSIL 1403

DB 1009 LTVNAQLREDANVCNKFHLDVSVENVQLNLKAKGAKALAKICTRYLSEVSTWII 1068
QY 1404 DISMWTGAPDDTDLKOLANGVDRIYSKYELOKAFSDORTLIIYLDKVSHEDEDDCLAFKV 1463
DB 1069 DVSMLTGFEVDTEDLTLSKGVDRY:SMFEINNNMAQKGTIIYLDKVSHEDEDECLAFKI 1128
QY 1464 HQYFNVELIOPGAVKYVAYYNLEESCTRFVHPHEKEDKLNKLCRDELRCRCAEENCFIOKS 1523
DB 1129 LKHEVGFIOGSKVYVYNLDEKCTKIYHDEATGLLNKICVGNVCRCAETCSLLNQ 1168
QY 1524 DDKVTLBERLDKACEPOVDYVYKTRLVKVLQSLNDFEYI:MAEQTIKSGSDEVQVQOQT 1583
DB 1189 QKNTROLRIQKAPDPNDVYVYKTLRLEEKDNDIYVMDVLEVLKQGTQDQOQVKVRQ 1248
QY 1584 FISPIKREALKLEEKHYLMWGLSSDFWGEKPN:SYIIGKDTWVHWPEDDECOBENQ 1643
DB 1249 YVSQKQCEALNLMVNDY:JWGSPSSDLWPMKDKISYLITKNTWIERPHEDKCOEERFQ 1308
QY 1644 KQCODLCAFTESM 1656
DB 1309 KLCDDFALFSYAM 1321

RESULT 15
US-09-925-442-34
Sequence 34, Application US/09925442
Patent No. 6607897
GENERAL INFORMATION:
APPLICANT: VOGEL, CARL-WILHELM
BREDEHORST, REINHORST
KOCK, MICHAEL
FRITZINGER, DAVID
TITLE OF INVENTION: RECOMBINANT PROCVF
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/925,442
FILING DATE: 10-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/017,947
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-0107-0X
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 1333 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-09-925-442-34

Query Match 41.1%; Score 3536; DB 4; Length 1333;

Best Local Similarity 51.2%, Pred. No. 0;
Matches 683; Conservative 254; Mismatches 372; Indels 24; Gaps 15;

334	QY	SGSDVVOAERSGIPVITSPYQIHTTKTKYFKPGMPEDLMVFTNPDGSPAYRPV---A	390
3	DB	SGGMMWTEQSGIHIVTSPYQIHTTKTKYFKPGMPYELTVYTKPDGSPAAPVPPVSEA	62
391	QY	VQGEDTVQSLTQSGGVAKLSINTHPSOKPISLITVTKKQELSEAEQATRMQALPYSTVG	450
63	DB	IHSEGT---TISDGTAKLFLNTPQNAQSLPITVTRTNHGDLPRERQAIKSMTATAYQTQ	118
451	QY	NSNNYHLVSLRTELRCETLNVNLFLLRMDRAEAKRYVYTYLIMNKGRLLKAGROVREP	510
119	DB	CGSNLHVATITSTIKPDGNLVPVNFNGVANSNLQIKYTYLILNKGKIFKVGROHRGD	178
511	QY	QODLVPLSTITDFIPSRFLVAVYTLIGASQREV-VADSVWVVDKSCVGLSVKSGQ	569
179	DB	C-NLVTNMLHTPDLIFSPREFVAYQV---GNNEIEVADSVWVVDKDTCMGLTVVKGAT	233
570	QY	SED-RQPVPGQOMTKIEGDHGAFLVAVDKGVFLNKKNKLTQSKIMDVWEKADIGCT	628
234	DB	SRDNRIQWGAAMKILEGDPGAWIGLVAVDKABYVLYNDKYKISQAKINWTIEKSDFGCT	293
629	QY	PGSGKDYAGVSDAGLITSSGGQOTQRAELQCPQPAARRR-SVCLTEKRMCKVKYK-686	
294	DB	AGSGONNIGVFEDAGLALTSTNLTNKTQSAKCPQANRRRRSSVLLDSNASKAAQFQ	353
687	QY	PKELRKCCEDGMRENPMRFSQORTRPISLGEACKVFLDCCNYITELRQOHARASHLGL	746
354	DB	QCDLAKCCEDGMHENPMGHTCERREKVTQEGDACKAAFLCCHYKIGIQDCKNRESELEFL	413
747	QY	ARSNLDEDIABENIVRSPEPESWLVNVEDLKEPPKN-GISTKMLNIFLKDSITITWEIL	805
414	DB	ARSDPEDDLFCBGHITSRSDPPESWLMLEQSHPNKSGISSKIVPFYLRDSITITWELL	473
806	QY	AVMSDDKKGICVADPFVTVMDPDLRLPLYSVVRNEQVEIRAVLYNQNOELKVRVE	865
474	DB	AVGSPTKGI CVABPEYETVMKDFIDQLPLYSVVRKNEQVKIRAVLYNY-ADKDIYVRVE	532
866	QY	LHNPAFCSLATRRRHQOTITIPKSSLSVPYVIVPLKTLQOEVEVKAAYVHHFISDGV	925
533	DB	LLYSPAFCSAGTESQRYEQLPIKALSRAVSFVIVPELQGLHDVEVTVASVQGELMSDGV	592
926	QY	RKSLKVVPEGIRMNKTAVARTLDPERLREGVQVKEDI PPADLSQVDPDTESETRILLQGT	985
593	DB	KKKLVVPEG-EWKSIVTIIELDPHTKIGIGTQVELYKANKLNDRVPDTEIEKIIITQGD	651
986	QY	PVAQWTEPAVDAERLNHILVTPSGCGGONMIGTPTVIAVHYLDETEQWEKFKLEKROGA	1045
652	DB	PVAQTIENSIDGSKNLHLIITPFGCGGONMIRAPVIAITYYLDITTOQWETLGINRTEA	711
1046	QY	LELIKKGTYQOLAFRQPSAPAAFKCAPSTWLTAYVYVVFSLAVNLIA-IDSQVLCGAV	1104
712	DB	VNQIMTGVAQQLVYKADHSYAAFTNSASSSWLTAYVYVKIFALAAKIVKDIHEIVCGGM	771
1105	QY	KWLILLEKOKPGVFOEDAPVTHQEWIGGLRNNNEKDMALTAFVLISLQEAKDICEQVNS	1164
772	DB	RALILNRTQOVRENAPVLFGTMQGGIQ-GAEPESGLTAFILVALLESRSICNAYINI	830
1165	QY	LPGSTTKAGDFLEANYMNLQRSYTVAIJAGYALAQMKLGKPLANKFLITTAOKNRWEDPG	1224
831	DB	LDSSISKATDYLLKKYKELQRPYTTALTAYALAAAERLNDRV--LMAASTGRNRWEEFN	888
1225	QY	KQLVNVEATSYALLALILQKDFDFVPPVVRWLNQRVYGGYGSTQATFMVFOALAQYOK	1284
889	DB	AHTHNIETSYALLALIKMKXFEAGPVVQWLIIDQYIYGGTYGQTQATVMVFOALAEYEI	948
1285	QY	DAPDHOELNLSQLPSRSSKITHRIHWESASLLRSEETKENEGFTVTAEGKQGTLSV	1344
949	DB	QNPETHKUNLIDITIELPREVPIRIRINENALAQTVETKLNEDFTVVSAGSGDKATMTI	1008
1345	QY	VTMYHAKA-QDLTCNKPDLKVTIKPAPETEKRPQDAKNTWILIEICTRYRGDQDATMSIL	1403

Search completed: December 22, 2004, 00:46:42
Job time : 53 secs

This Page Blank (uspto)